MICROBIOLOGICAL REVIEWS, Dec. 1985, p. 379–397
0146-0749/85/040379-19$02.00/0

Vol. 49, No. 4

# Structure and Organization of the Transfer Ribonucleic Acid Genes of *Escherichia coli* K-12

MAURILLE J. FOURNIER[1]* AND HARUO OZEKI[2]

*Department of Biochemistry, University of Massachusetts, Amherst, Massachusetts 01003,[1] and Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan[2]*



## INTRODUCTION

The structural genes for transfer ribonucleic acid (tRNA) constitute a particularly unique class of genetic material. The coding regions corresponding to the mature RNA product are small and more or less uniform in size, ranging from about 75 to 90 base pairs (bp). Despite this small size there is an amazing amount of structural and functional information encoded in each transfer deoxyribonucleic acid (tDNA) element (tDNA is the DNA corresponding to tRNA). The nucleotide sequences and biological roles vary markedly for different tRNAs, yet all assume an essentially common structure in solution. The common secondary and tertiary features are determined by regularly spaced stem-loop elements and other combinations of positionally fixed complementing nucleotides. At the DNA level there are four and sometimes five inverted repeats which define the familiar cloverleaf, and up to 18 positionally constant nucleotides which dictate the tertiary configuration. In addition to these elements each gene also encodes the common CCA (C, cytosine; A, adenine) sequence which occurs at the 3′ terminus of all tRNAs. Together, the nucleotides specifying these structural features of the product utilize nearly 80% of the bases in the corresponding tDNA sequence.

On the surface then, it would appear that most of the informational capacity of the tRNA gene is consumed just in specifying the basic structural requirements of the final product. We know that this is not the case, of course, and that the tDNA elements contain an extraordinary amount of functional information as well. This information is both integrated with and superimposed on the sequence elements that define the structure. Both unique and common functional properties are encoded in each gene. Included are recognition sites for a considerable array of proteins involved in precursor maturation, protein biosynthesis, and a variety of nontranslational processes. In *Escherichia coli*, a minimum of 8 to 10 enzymes are involved in just converting the primary transcript to the mature tRNA molecule. This group includes the nucleolytic processing activities which trim away extra nucleotides from the ends and upwards of six unique base-modifying enzymes (7, 30, 70).

When the proteins involved in the functioning of the mature tRNA are considered, the number of specific recognition elements encoded in the tRNA gene becomes even more impressive. tRNAs participating in translation must be recognized by a specific aminoacyl-tRNA synthetase, initi-

* Corresponding author.

ation, elongation, and termination factors, and components of the ribosome (106). tRNAs with other functions will interact with yet other proteins. In this last regard, it has been known for some time that *E. coli* tRNAs can serve as amino acid donors in the secondary modification of proteins (89), as regulators of transcriptional termination by attenuation—in the cases of certain of the amino acid biosynthetic operons (102, 105), and as a regulator of branched-chain amino acid transport (74); additional functions have been documented for other procaryotic organisms and eucaryotic and viral sources (reference 17 and references therein). New additions to the growing list of functions encoded in the *E. coli* tRNAs come from recent discoveries identified below that reveal that one *E. coli* tRNA may be required for maintenance of normal DNA replication activity (66; S. Nishimura, personal communication) and another is required for completion of the cell division cycle (94).

Thus, when taken together, the total number of unique proteins with which a single tRNA interacts during its lifetime is probably in excess of 20. While there was a time when some wondered why tRNAs were so large, acknowledging only the need for anticodon and amino acid attachment sites, the general view now must be one of outright wonderment at the remarkable amount of information concentrated in these small molecules and, therefore, their genes.

In *E. coli* a feature that distinguishes the tRNA genes from those coding for proteins is the frequent occurrence of multiple gene copies. Redundancy is a property of the ribosomal RNA (rRNA) genes as well, where the genomic copy number is seven (23). As we shall detail below, the complexity of the tRNA gene sets is enhanced even further by the knowledge that these elements occur in three different types of transcription units, specifying (i) tRNA only, in both homologous and heterologous gene sets, (ii) tRNA and the three rRNAs, and (iii) tRNA and protein.

The development of gene-cloning techniques and rapid DNA sequencing methods has, of course, completely revolutionized our approaches to describing gene structure and function. In the case of the tRNA genes the application of these technologies has recently yielded a considerable amount of new and valuable information. The impact of these methods is reflected in the rate at which new gene sequences are being developed. Before 1980, there was only partial DNA sequence information available for two classes of *E. coli* tRNA gene systems. These were the *tyrT* operon with tandemly duplicated tyrosine tRNA ($tRNA_1^{Tyr}$; subscript refers to the isoacceptor species) genes, characterized as the site of the $Su_3^+$ mutation, and segments of three rRNA operons with isoleucine tRNA ($tRNA_1^{Ile}$) and alanine tRNA ($tRNA_{1B}^{Ala}$) in the 16S to 23S spacer region (22, 86, 112) and aspartate tRNA ($tRNA_1^{Asp}$) and tryptophan tRNA ($tRNA^{Trp}$) in a region following the 5S RNA gene (110). At the present time, only 5 years hence, over 40 *E. coli* tRNA genes have been sequenced (the literature survey for this review was concluded in June 1985).

The 1983 edition of the linkage map of *E. coli* K-12 identifies the chromosomal locations of about 1,000 genes (5). Among these, 45 correspond to genes coding for 28 different tRNA subspecies. The map locations of 9 additional tRNA genes have since been reported, bringing the current total number of assigned genes to 54 and the number of unique subspecies encoded to 35 (see below). This last value compares with a cellular content of unique tRNA species estimated to be about 60, based on inspection of fractionation patterns achieved by column chromatography and two-dimensional electrophoretic methods (D. A. McKay and M. J. Fournier, unpublished observations). This estimate could be somewhat low considering that some cofractionation of species can occur and that species in low relative abundance can escape detection. Offsetting this to some extent are possible overestimates arising from the occurrence of separable but genetically identical isoacceptors that differ only in modified base content. The total number of tRNA genes encoded in the chromosome is not known precisely but has been roughly estimated by hybridization assays to be about 60 (11). Extrapolating from the 30% redundancy level known to exist for the nonribosomal tRNA cistrons (see below), 50% for all tRNA genes known, the total complement of tRNA genes could be about 75. Thus, with over 50 genes mapped, it appears that at least half and perhaps as many as two-thirds of the actual total have already been detected.

In this review we summarize the current state of knowledge about the structure and organization of the tRNA genes of *E. coli* K-12. Special emphasis is placed on the genes which have been mapped and sequenced. The results reviewed are confined to the bacterium-specific tRNA genes; genes encoded in the genomes of the coliphage are not included. Although the focus is on the genes of *E. coli*, we do comment on the tRNA genes from two other procaryotic organisms as well. We include information about the tRNA genes of *Salmonella typhimurium* which show that the strong genetic resemblance this organism has with *E. coli* also extends to its tRNA genes. Finally, because this review follows one devoted to the tRNA genes of *Bacillus subtilis* (103), we also compare the genes of *E. coli* and *B. subtilis*. This review is the first since 1980 on the genetic organization of the *E. coli* tRNA genes (72, 73) and the first ever to comparatively assess these genes at the DNA level. (A compilation of the DNA sequences for the nonribosomal tRNA genes has been prepared and will be made available on request.)

## GENE MAPPING AND GENOMIC ORGANIZATION

A mix of direct and indirect approaches have been used to determine the chromosomal locations of the tRNA genes. While many have been mapped in studies directed to that end, the positions of several others have become known as unexpected bonuses in the characterization of neighboring genes or of tRNA functions not previously recognized as such. We review here the approaches and results from the various mapping studies completed to date. References to chromosomal mapping are provided only for results developed since the 1983 edition of the complete linkage map (5). Other mapping strategies that will likely be used in the future are identified in the Discussion section. The current results are summarized in the tRNA gene linkage map shown in Fig. 1 and in Table 1.

### Mapping of Suppressor tRNA Genes

Many tRNA cistrons have been mapped by locating suppressor genes that were later determined to be tRNA structural gene mutants, most with altered anticodon sequences (73). This approach was the first used and still ranks among the most direct and convenient means of mapping available today. Before the advent of modern cloning technology, most of the detailed information about tRNA gene structure

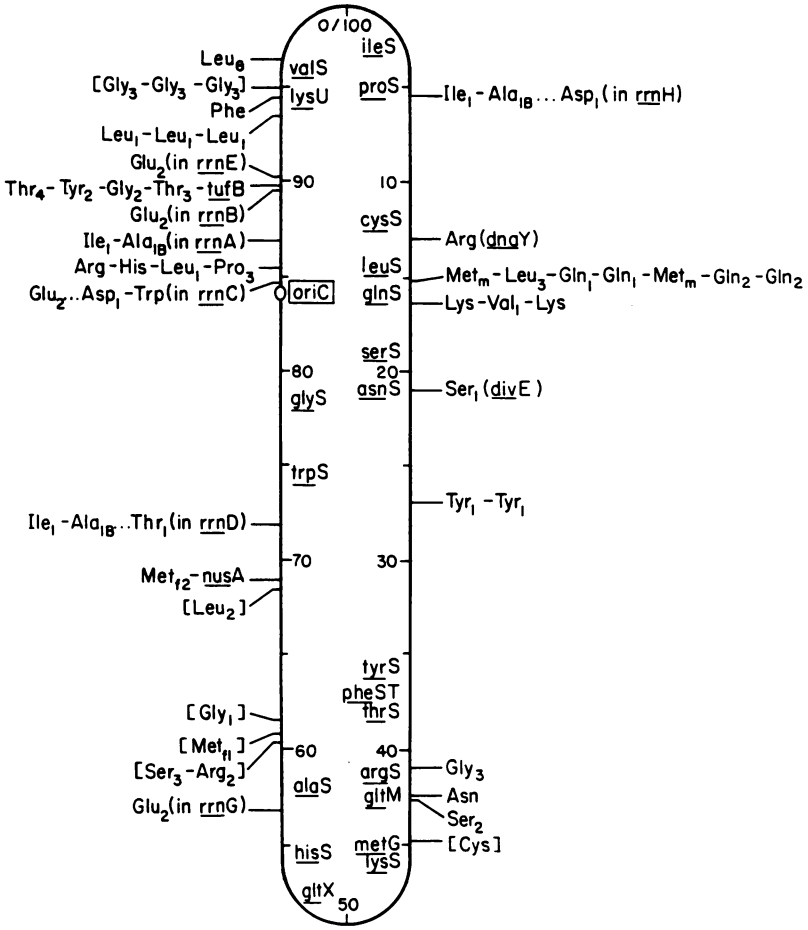



FIG. 1. Genetic map locations of the *E. coli* genes for tRNA and aminoacyl-tRNA synthetase. The identities of the tRNA gene regions appear outside of the chromosomal map, while those for the synthetases are inside. The numbers correspond to minutes on the linkage map, and the position of the origin (*oriC*) is shown (5). The tRNA gene designations specify the various isoacceptor species encoded and order of transcription (5′ → 3′). Brackets identify genes for which DNA sequence information is not yet available. The synthetases are derived from single-copy genes except for the lysine-specific activity which is encoded at two sites, *lysS* and *lysU*; the latter is expressed upon heat shock (104). The genetic nomenclature for the various genes is given in Table 1 along with map positions and alternate designations.

and function came from studies of transducing phages carrying various nonsense suppressors (72, 73).

To date, 10 chromosomal sites coding for 11 suppressor tRNA genes have been located by direct mapping analyses. As will be detailed below, many suppressor genes occur with other tDNA sequences, both related and unrelated; however, some also stand alone. Recently characterized members of this last class are genes for suppressor variants of glycine tRNA ($tRNA_3^{Gly}$) (*SuA58-SuA78* [101]), serine tRNA ($tRNA_2^{Ser}$) ($Su_1^+$; [92, 93, 96; M. Yoshimura et al., manuscript in preparation]), and leucine tRNA ($tRNA_6^{Leu}$) ($Su_6^+$ [95, 109]). As luck would have it, sequence analysis of the $Su_1^+$ DNA revealed the presence of a second tRNA gene about 1 kilobase away. The bonus gene appears to define a transcription unit encoding asparagine tRNA ($tRNA^{Asn}$) (Yoshimura et al., in preparation).

### Mapping of Nonsuppressor Functions

The chromosomal positions of three tRNA genes have been recently assigned through mapping of functions unrelated to tRNA suppression. In two instances mapping involved placement of complementing functions previously unknown to be associated with tRNA.

DNA clones corresponding to phenylalanine tRNA ($tRNA^{Phe}$) genes have been selected and mapped by complementation of a temperature-sensitive mutation in a subunit of the cognate aminoacyl-tRNA synthetase (15, 28). The success of this approach presumably reflects protection of the synthetase against thermal inactivation by $tRNA^{Phe}$ overproduction.

The two tRNA genes discovered and mapped through analysis of apparently unrelated functions correspond to loci associated with cell division (94) and DNA synthesis (66). The discovery that these loci encode tRNAs came as a complete surprise as tRNA had not been directly implicated in either process before. The mutation affecting cell division is a temperature-sensitive allele of the *divE* locus, while that affecting replication is known as *dnaY*. Sequencing of DNA from the *divE* locus showed it to correspond to $tRNA_1^{Ser}$ (94). Interestingly, the basis of the temperature-sensitive effect is an A to G (G, guanine) change in the middle of the stem to the dihydrouridine loop. Sequence and expression analyses with deletion derivatives of the *dnaY* gene revealed it to encode a small RNA (66). The sequence was subsequently recognized to correspond to a previously uncharacterized

TABLE 1. tRNA gene symbols

| tRNA | Anticodon | Gene symbol[a] | Map position[a] (min) | Alternate symbols and remarks | References |
|---|---|---|---|---|---|
| Ala 1B | UGC | *alaT* | 86.5 | In the *rrnA* operon | 23 |
| Ala 1B | UGC | *alaU* | 72 | in the *rrnD* operon | 23 |
| Ala 1B | UGC | *alaV* | 5 | In the *rrnH* operon | 23 |
| Arg | CCG | *argT* | (85) | | 37 |
| Arg | UCU | *argU*[b] | 12 | *dnaY* | 66 (see text) |
| Arg 2 | $^{I}_{A}$CG | *argV*[b] | (61) | Cotranscribed with *serV* | 44, 46, 81 |
| Asn | GUU | *asnT* | 43 | Cloned with *serU* | Yoshimura et al., in preparation; T. Sekiya, unpublished data |
| Asp 1 | GUC | *aspT* | 84.5 | In the *rrnC* operon | 23 |
| Asp 1 | GUC | *aspU* | 5 | In the *rrnH* operon | 23 |
| Cys | GCA | *cysT* | (45) | | 44; Ikemura, personal communication (see text) |
| Gln 1 | UUG | *glnUα*[b] *glnUβ*[b] | 15.5 | *supB SuB*; tandemly duplicated gene | 67 |
| Gln 2 | CUG | *glnVα*[b] *glnVβ*[b] | 15.5 | *supE Su2*; tandemly duplicated gene | 48, 49, 67 |
| Glu 2 | UUC | *gltT* | 90 | In the *rrnB* operon | 23 |
| Glu 2 | UUC | *gltU* | 84.5 | In the *rrnC* operon | 23 |
| Glu 2 | UUC | *gltV* | 90.5 | In the *rrnE* operon | 23 |
| Glu 2 | UUC | *gltW* | 57 | In the *rrnG* operon | 23 |
| Gly 1 | CCC | *glyU* | 62 | *supT sufD SuA36* | 24, 35 |
| Gly 2 | UCC | *glyT* | 90 | *SuA36* | 24 |
| Gly 3 | GCC | *glyW* | 42 | *SuA58 SuA78* | 101 |
| Gly 3 | GCC | *glyVα*[b] *glyVβ*[b] *glyVγ*[b] | 95 | *SuA58 SuA78*; tandemly triplicated gene | 47, 81 |
| His | GUG | *hisR* | (85) | | 37 |
| Ile | GAU | *ileT* | 86.5 | In the *rrnA* operon | 23 |
| Ile | GAU | *ileU* | 72 | In the *rrnD* operon | 23 |
| Ile | GAU | *ileV* | 5 | In the *rrnH* operon | 23 |
| Leu 1 | CAG | *leuT* | (85) | | 41, 44 |
| Leu 1 | CAG | *leuVα*[b] *leuVβ*[b] *leuVγ*[b] | (93) | Tandemly triplicated gene | 19, 44 |
| Leu 2 | GAG | *leuU* | (68) | | 44 |
| Leu 3 | UAG | *leuW* | 15.5 | In the *supB-E* operon | 67 |
| Leu 6 | CAA | *leuX* | 97 | *supP Su6* | 95, 108 |
| Lys | UUU | *lysTα*[b] *lysTβ*[b] | 16.5 | *supG supL Suβ*; duplicated in a tRNA operon with *valT* | 109 |
| Met m | CAU | *metTα*[b] *metTβ*[b] | 15.5 | Duplicated in the tRNA operon *supB-E* | 49, 67 |
| Met f1 | CAU | *metZ* | (61) | | 44 |
| Met f2 | CAU | *metY* | 69 | | 44, 51 |
| Phe | GAA | *pheU* | 94.5 | | 28, 83 |
| Pro 1 | CGG | *proV*[b] | Unknown | Sequenced but unmapped | Kuchino and Nishimura, in preparation |
| Pro 3 | UGG | *proU*[b] | (85) | | 37 |
| Ser 1 | UGA | *serT* | 22 | *divE* | 94 |

*Continued on following page*

TABLE 1—*Continued*

| tRNA | Anticodon | Gene symbol[a] | Map position[a] (min) | Alternate symbols and remarks | References |
|---|---|---|---|---|---|
| Ser 2 | CGA | *serU* | 43 | *supD supH Sul* | 92, 96 |
| Ser 3 | GCU | *serV* | (61) | | 44, 46, 81 |
| Thr 1 | GGU | *thrV* | 72 | In the *rrnD* operon | 23 |
| Thr 3 | GGU | *thrT* | 90 | | 4, 78 |
| Thr 4 | UGU | *thrU* | 90 | | 4, 78 |
| Trp | CCA | *trpT* | 84.5 | *Su7 Su8 Su9*; in the *rrnC* operon | 23, 110 |
| Tyr 1 | GUA | *tyrT* | 27 | *supF Su3 Su4 SuC*; tandemly duplicated with *tyrV* | 22, 78, 79 |
| Tyr 1 | GUA | *tyrV* | 27 | *supF Su3 Su4 SuC*; tandemly duplicated with *tyrT* | 22, 78, 79 |
| Tyr 2 | GUA | *tyrU* | 90 | *supM supB15* | 4, 40, 78 |
| Val 1 | UAC | *valT* | 16.5 | Contiguous with *lysT* | 109 |

[a] From the linkage map of *E. coli* K-12, edition 7 (5). Map positions in brackets are approximate.
[b] Gene symbols proposed in this paper.

arginine tRNA ($tRNA^{Arg}$) with a UCU (U, uracil) anticodon (Nishimura, personal communication). Recombinants carrying the cloned *dnaY* region have since been shown to overproduce arginine tRNA (J. R. Walker, personal communication). It remains to be determined whether this locus has a direct role in the replication process.

### Mapping with Episomes and by Prophage Induction

Another mapping strategy that has been used with considerable success is based on correlating increased production of specific tRNAs with amplification of known regions of the chromosome. This has been achieved by using merodiploids harboring specific F′ factors and by inducing prophages of known chromosomal location (44). The mapping results obtained with F′ factors are naturally of lower resolution owing to the appreciable size of the DNA region amplified in each case. Of course, mapping to higher resolution can be accomplished by extending the analysis with hybrid plasmids or phage derivatives carrying smaller portions of the appropriate genomic segment. These two approaches to mapping can be used to identify and place individual coding sequences and also for general screening of yet undefined tRNA genes. A hypothetical source of confusion in both cases is the possibility that enhanced production of tRNA could in some instances reflect a change in expression that is unrelated to a simple increase in tRNA copy number. The fact that this situation has not arisen yet suggests that it may only be a theoretical risk.

Altogether, assignments for 26 tRNA genes were determined in a single study by using these strategies (44); the F′ mapping yielded the gross locations of 14 genes, while sites for 24 tDNAs were positioned within 1 to 2 min in the prophage induction analysis. Most of the genes initially mapped in this way corresponded to tRNAs of known identity; however, some did not. Identification of the RNA species overproduced was by reference to two-dimensional maps of electrophoretically separated tRNAs and catalogs of RNA fingerprints. This means of identification is limited, of course, to RNA species which have already been at least partially characterized. Five of the unidentified tRNAs mapped in the early study and given letter designations have now been identified. These species are: E = $tRNA_6^{Leu}$ (95, 108), M = $tRNA_2^{Ser}$ (92, 93, 96), P = proline tRNA ($tRNA_3^{Pro}$) (37), X = $tRNA^{Leu}$ with a UAG anticodon (67) which is designated $tRNA_3^{Leu}$ (72), and C = cysteine tRNA ($tRNA^{Cys}$) (T. Ikemura, personal communication).

### tRNA Gene Clusters

The chromosomal locations of several tRNA genes have become known through sequence characterization of tDNA clusters that include at least one previously mapped gene. Linkage of different mapped and unmapped tDNAs has been revealed by sequencing of the DNA itself or an RNA transcript, the latter corresponding to a multimeric precursor. Gene clusters defined in this way include $Su^+$ genes isolated in transducing phage and $Su^-$ gene sets cloned by in vitro methods. Sequencing of $Su^+$ DNA has shown that suppressor genes can occur in clusters with a corresponding $Su^-$ allele or with tRNA genes unrelated to the suppressor. $Su^+$-containing operons characterized in this fashion include those encoding the tRNA species $Tyr_1$-$Tyr_1$ (3, 22, 62, 77, 79, 84), $Met_m$-$Leu_3$-$Gln_1$-$Gln_1$-$Met_m$-$Gln_2$-$Gln_2$ (67), $Thr_4$-$Tyr_2$-$Gly_2$-$Thr_3$-*tufB* (4, 78), and Lys-$Val_1$-Lys (108) (Gln, glutamine; Thr, threonine; Lys, lysine; Val, valine). The structures of these and the other tRNA transcription units characterized to date are shown in Fig. 2.

Defined in similar fashion but without benefit of $Su^+$ selection are tRNA operons for Arg-histidine (His)-$Leu_1$-$Pro_3$ (37) and $Leu_1$-$Leu_1$-$Leu_1$ (19). These gene sets were isolated by direct cloning and hybridization and located on the chromosome with the aid of previous F′ and prophage mapping data. The structure of a yet unmapped genomic segment containing a $tRNA_1^{Pro}$ gene has also been determined in this manner (54a). Information about the structures and map locations of several gene sets has come from characterization of multimeric tRNA precursors that accu-

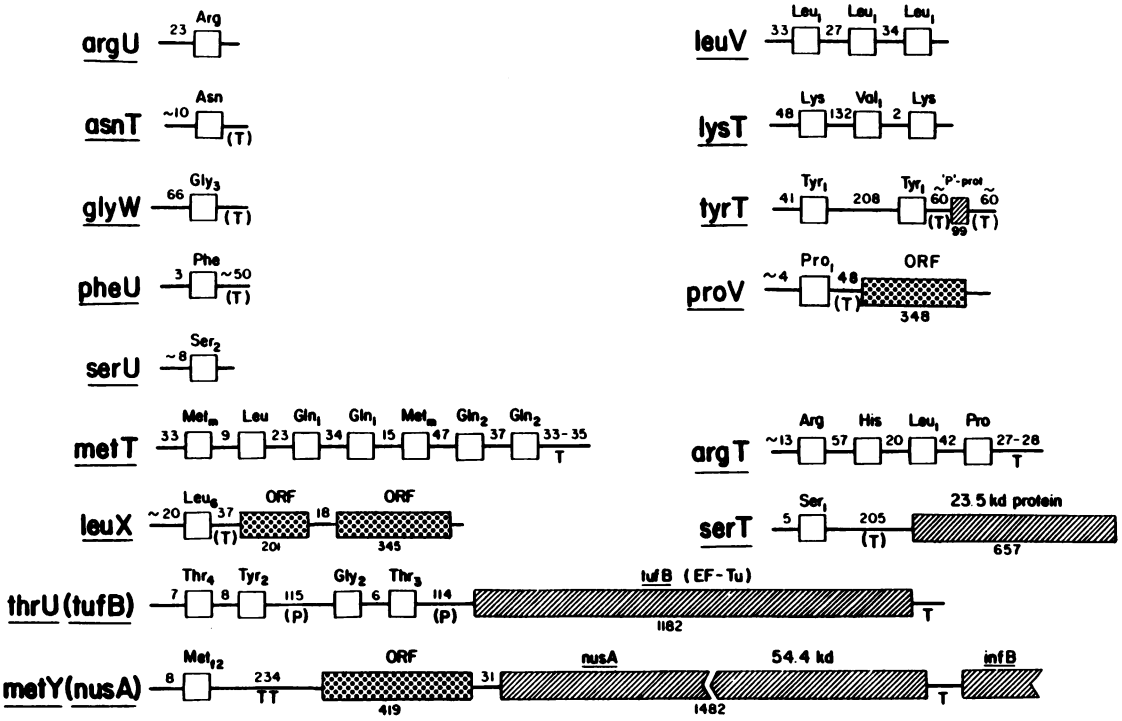



FIG. 2. Structures of the sequenced tRNA gene regions of *E. coli* K-12. The organizations of 15 genomic tDNA segments for which sequence information is available are shown. Each gene region is identified by its genetic designation. The structures are arranged with the known or likely promoter elements at the left end. Individual tRNA genes are represented by open boxes, and the identities of the acceptor species encoded are given above the respective genes. Protein genes and extended ORF sequences are represented by hatched boxes, and sizes (in base pairs) of the coding regions are shown below the boxes. The symbol T identifies transcriptional terminators shown to be functional in expression analyses, while candidate terminators identified solely by sequence inspection are designated by the symbol (T). Candidate promoter elements in the *thrU* operon are labeled (P). The numbers flanking the individual coding regions correspond to the number of nucleotides separating that gene from its promoter, a neighboring gene, or a demonstrated terminator element.

mulate in processing-defective cells. The precursors from the ribonuclease P conditional mutants have been especially valuable in this regard (30, 46, 81, 87). As before, knowledge about the map position of one gene in the cluster allowed ready placement of the entire set. The tRNA operons encoding the tandemly triplicated $tRNA_3^{Gly}$ subspecies (81) and the linked $Ser_3$-$Arg_2$ tRNAs (46, 81) were detected and mapped by this strategy. A gene cluster encoding the isoacceptors $tRNA_{2A}^{Val}$ and $tRNA_{2B}^{Val}$ has been identified through precursor sequencing but remains unmapped owing to the absence of data for any adjoining gene (46, 81).

## rRNA Operons

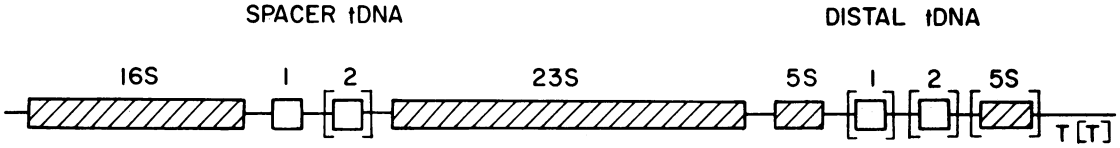
Principal among the tDNAs that have been detected and mapped through the characterization of unrelated genes are those encoded in the rRNA operons. The complex nature of these operons was initially established from sequence studies of rRNA precursors and the rRNA genes rather than through characterization of the resident tRNA cistrons (23, 30, 60, 65). Similarly, mapping of the individual *rrn* operons was based on rRNA gene function rather than tDNA activity.

Each of the seven rRNA operons in *E. coli* has been shown to contain at least one tRNA gene, two operons are known to encode two tRNAs, and others include three tDNAs (23, 65). All of the *rrn* operons contain tDNA in the spacer region between the 16S and 23S rRNA genes; some also contain tRNA cistrons at the distal end of the transcription unit. The 16S to 23S spacer genes code exclusively for glutamate tRNA ($tRNA_2^{Glu}$) or a dimeric precursor to the $Ile_1$ and $Ala_{1B}$ tRNAs. The distal tDNAs encode $Asp_1$, $Thr_1$, or Trp tRNAs. Since the map locations of each of the individual rRNA operons is known, the positions of the associated tDNA sequences have also been established. The map locations of the rRNA operons are also shown in Fig. 1, and the organization of the individual transcription units is depicted in Fig. 3.

## Clustered tRNA and Protein Genes

Results from expression and DNA sequence analyses have revealed that at least six tRNA cistrons occur in tandem with known or potential protein-coding regions. Genetic map locations have been determined for five of the six loci. Two of these clusters have been demonstrated to be true mixed-function operons; for the others it remains to be determined whether the adjacent open reading frame (ORF) sequences are actually expressed.

The two cases in which coexpression has been demonstrated both in vivo and in vitro are the operons encoding (i) the four tRNA species $Thr_4$-$Thr_2$-$Gly_2$-$Thr_3$ and elongation factor Tu (40, 59, 64) and (ii) the *tyrT* operon which contains duplicated genes for $tRNA_1^{Tyr}$ and a cistron for a small, arginine-rich P protein of yet unknown function (2, 31, 55, 56, 77). The sequence and expression results obtained for these and the other tRNA operons are reviewed below. The four cases in which a tRNA coding region has been shown to




| rRNA Operon | Chromosomal location (min) | Spacer tDNA 1 | Spacer tDNA 2 | Distal tDNA 1 | Distal tDNA 2 |
|---|---|---|---|---|---|
| rrnA | 86.5 | $Ile_1$ | $Ala_{1B}$ | | |
| rrnB | 90 | $Glu_2^{\oplus}$ | | | |
| rrnC | 84.5 | $Glu_2$ | | $Asp_1^{\oplus}$ | $Trp^{\oplus}$ |
| rrnD | 72 | $Ile_1^{\oplus}$ | $Ala_{1B}^{\oplus}$ | $Thr_1^{\oplus}$ | |
| rrnE | 90.5 | $Glu_2^{\oplus}$ | | | |
| rrnG | 57 | $Glu_2$ | | | |
| rrnH | 5 | $Ile_1$ | $Ala_{1B}$ | $Asp_1^{\oplus}$ | |

FIG. 3. Structures of the rRNA operons of *E. coli* K-12. The overall organization of the seven *rrn* operons is shown in schematic form in the upper part of the figure. Brackets identify tRNA or 5S RNA coding regions that are not common to all operons. Transcription direction is left to right terminating in a region containing one or two terminator elements, T[T]. The identities of the known tRNA genes are given for each operon in the lower portion of the figure. Sequenced tDNA regions are identified with the symbol ⊕; references are provided in the text. The *rrnD* operon contains a second gene for 5S RNA, following the distal $tRNA_1^{Thr}$ gene.

be in tandem with one or more ORF elements correspond to the genes encoding tRNA subspecies $tRNA_6^{Leu}$ (95, 108), $tRNA_{f2}^{Met}$ (50, 51), $tRNA_1^{Ser}$ (94), and $tRNA_1^{Pro}$ (Kuchino and Nishimura, in press). The composition of these gene sets is shown in Fig. 2, and the genomic locations are provided in Table 1.

### Chromosomal Distribution of the tRNA Genes

Altogether, the linkage map in Fig. 1 shows the chromosomal locations of 54 tRNA cistrons encoding 35 different tRNAs. The genes for 19 of the 20 aminoacyl-tRNA synthetases are also shown. Only the gene for aspartyl-tRNA synthetase remains to be mapped. The mapping results show both the tRNA and synthetase genes to be dispersed throughout the chromosome with no correlation existing between the genomic positions of the sequences coding for synthetase and those of the corresponding tRNA. Similarly, except for tandem repeats of certain tDNAs there is no obvious pattern to the chromosomal distribution of the individual tRNA genes. A number of genes are concentrated near the origin of replication, especially those that are integrated with the *rrn* operons. However, the majority of tRNA genes are located elsewhere in the chromosome. The location near *oriC* does offer the potential of greater overall expression due to the higher gene copy number that occurs in periods of rapid growth. This advantage can be appreciated for the rRNA genes but is not apparent for the particular tRNA genes in this region.

While the map shows the tRNA genes to be widely distributed, there is at the same time a considerable degree of clustering of the individual cistrons. Of the 54 mapped genes, 37 (68%) are organized in polycistronic operons containing other tRNA or rRNA genes. Seven of these clusters correspond to the rRNA operons which encode one to three tRNAs each and together account for a minimum total of 14 tRNAs. Even if the tRNA cistrons of the *rrn* operons are considered separately, 30 of the 40 remaining genes (75%) mapped thus far are known to occur in transcription units with two or more tRNA cistrons. Except for the iteration of identical and closely related tDNAs, the manner in which the individual tRNA genes are organized in the various operons does not suggest any obvious functional basis.

The now well-documented extensive clustering of the *E. coli* tRNA genes was first detected in characterization of bulk tRNA-DNA hybrids, well before detailed mapping and sequence results became available (11, 25). Density profiles of hybrids formed with purified tDNA fragments of approximately 400 bp in length revealed that most of the tRNA genes were clustered.

A higher-resolution study involving Southern hybridization assays has recently been completed which extends this earlier conclusion (16). In this latter investigation various restriction digests of chromosomal DNA were probed with bulk tRNA. The simple patterns resulting verified the widespread occurrence of clustering and further indicated that the total number of genomic segments containing tDNA may be limited to the order of 13 to 16; of these, 11 are distinct from the regions encoding rDNA. Taken with the current linkage map, these results suggest the possibility that many, perhaps even most, of the estimated 20 or 30 tRNA genes yet to be identified will eventually be placed in regions already known to contain tDNA.

Many of the tDNA genes are present in multiple copies, ranging in number from two to four. Multiple copies occur both in the same and different operons and in operons encoding tRNA only, tRNA and the rRNAs, and tRNA and protein. From the data available on tRNA and tRNA gene content, it appears that gene copy number plays a major role

in assuring a balanced population of acceptor species. A strong correlation exists between codon demand and tRNA content, and the relationship between tRNA abundance and gene copy number appears to show good correspondence as well (32, 41, 42, 45, 106).

It has been generally assumed that the occurrence of tRNA genes in the *rrn* operons somehow assures the coordination of ribosomal and bulk tRNA synthesis. Recent evidence for feedback regulation of rRNA operons has led to a proposal of a common negative-control system for both rRNA and tRNA transcription, but a firm connection has yet to be made (52, 71). The identities of the tRNAs encoded in the *rrn* operons does not suggest anything special in this regard. The basis for their occurrence and maintenance in the *rrn* operons constitutes one of the more interesting questions about tRNA gene organization and expression.

Also not understood at this time are the bases for the organization of the non-*rrn* tRNA operons. This population of operons includes transcription units with one to seven tRNA cistrons and gene systems that include protein-coding regions as well as tRNA genes. As shown in the schematic representations in Fig. 2, four of the six operons encoding multiple tRNA genes contain repeats of identical or closely related tDNA. The presence of such repeats in tandem or near-tandem array presumably reflects the occurrence of gene duplication events, likely arising through nonreciprocal recombination. The occurrence of sequence-related tDNAs at different loci argues that duplication with translocation has also occurred. This arrangement, which obviously holds for the iterated *rrn* operons, also occurs for the separated operons coding for identical and closely related tRNA subspecies. Included are the $tRNA^{Arg}$ isoacceptors encoded by *argU* and *argT* (37, 66); $tRNA_1^{Gly}$ and $tRNA_2^{Gly}$; $tRNA_1^{Leu}$; $tRNA_{f2}^{Met}$ and $tRNA_m^{Met}$; $tRNA_1^{Pro}$ and $tRNA_3^{Pro}$; $tRNA_1^{Ser}$ and $tRNA_2^{Ser}$; $tRNA_3^{Thr}$ and $tRNA_4^{Thr}$; and $tRNA_1^{Tyr}$ and $tRNA_2^{Tyr}$ (90, 91).

### A Recommendation about Genetic Nomenclature

Questions about genetic nomenclature arise when the polycistronic transcription units are considered. First, is the issue of how to refer to the operon itself? Should the operon carry the name of one or several genes? If only one, should it be the gene which historically first identified the locus or simply the first gene in the sequence? A second question relates to the designations to be used for multiple copies of a tRNA gene that occur in the same transcription unit.

In the cases of transcription units that encode only tRNA it will be relatively easy to adopt the convention of naming the operon for the first gene. The only difficulty here is the historical precedent of referring to some of these loci with designations now known to correspond to genes that are internal in the operon, e.g., the *supB-E* locus. While we do not propose that these familiar designations be dropped, we do recommend adopting the convention of naming the operons for the first gene of the transcription unit and including this usage in all reports. Thus, the operon encoding the *supB-E* genes will carry the designation *metT*.

Identifying nomenclature that is appropriate for the mixed transcription units poses no problem in the cases of the rRNA operons but is difficult for some of the jointly expressed genes encoding both tRNA and protein. The mixed tRNA-rRNA operons have traditionally only been referred to as *rrn* operons, and there does not seem to be a case to offer for renaming these gene sets. However, no convention currently exists for naming the apparently growing class of mixed tRNA and protein operons. The situation for this last class is complicated by the current use of now-familiar designations for genes that are not first in the transcription unit (e.g., *tufB*) and existing uncertainty about the biological functions of some of the encoded products. Naming an operon according to a major encoded function has a natural basis and obvious appeal. There is, of course, the question of how to proceed when multiple unique functions are encoded. Once again, simple designation of the operon by the first gene would alleviate the need for such choices. We recommend that this practice be used here as well as for the tRNA-only operons.

The existing system of nomenclature (5, 18) which uses letters of the alphabet to identify individual tRNA genes works well overall but does have limitations for repeated genes at the same locus. We propose here a refinement of the current system to address this limitation. With the endorsement of the Curator and Advisory Committee to the *Escherichia coli* Genetic Stock Center (B. Bachmann, A. J. Clark, K. B. Low, K. E. Sanderson, and A. L. Taylor), we recommend the use of lowercase Greek letters to distinguish the individual repeated cistrons at a common locus. Thus, for example, the three $tRNA_1^{Leu}$ genes of the *leuV* locus would be designated *leuV*α, *leuV*β, and *leuV*γ. Similarly, the duplicated pair of genes for $tRNA_1^{Gln}$ and $tRNA_2^{Gln}$ in the *metT* operon would be *glnU*α, *glnU*β, *glnV*α, and *glnV*β. This convention would also apply for noncontiguous repeats in the same operon. Thus, the nontandem repeated genes for $tRNA^{Lys}$ in the *lysT* operon would be designated *lysT*α and *lysT*β. We recommend that this system be adopted in all cases in which a letter designation is not already in use. There is now only one case in which the modified nomenclature would not be applied, that of the *tyrT* operon. In this case the repeated $tRNA_1^{Tyr}$ genes would continue to be designated *tyrT* and *tyrV* (5).

## ORGANIZATION OF THE tRNA TRANSCRIPTION UNITS

### tRNA Genes Occur in both Simple and Complex Transcription Units

At present, there is detailed DNA sequence information for 15 different tRNA gene clones encoding one or more tRNAs. The sequences for at least five of the tDNA-containing regions in the seven rRNA operons have also been determined. As already noted, considerable insight about the functional organization of many of these cloned genes has been developed from genetic analyses, various expression studies, and structural analyses of precursor tRNAs. From these results it has been established that the various gene systems fall into different functional classes. There are single-function transcription units that encode from one to seven tRNAs and multifunctional sets that direct the synthesis of tRNA and rRNA or tRNA and protein. Owing to the absence of results from expression studies, identification of the operon class to which the various nonribosomal tRNA genes belong can be considered firm in only a few cases. Despite this limitation the DNA sequence information does allow tentative classifications to be made. To the extent that recent history has shown that authentic promoter and terminator elements can be recognized by sequence inspection alone, it is reasonable to suppose that most of the suggested categorizations will probably be sustained. It is important to remember, of course, that not all

promoter elements resemble the canonical consensus sequence (34, 76) and only terminator elements of the rho-independent class can be recognized at this time (36, 76). Conversely, homology with a consensus sequence does not at all assure that an element has any functional relevance (34, 76).

With this caveat firmly in mind the population of cloned tDNA regions can be tentatively cataloged in the following way. Nine or more of the 15 non-*rrn* genomic segments may correspond to transcription units encoding tRNA only. The remaining six are known or have the sequence potential to define operons that also encode protein. Based on the DNA sequence results at least three tDNA regions appear likely to constitute monocistronic transcription units. Most, and possibly all, of the remaining gene regions could define multigene operons. The number of tDNA elements in the known tRNA operons varies from three to seven, while the known or suspected mixed function operons contain from one to four tRNA coding regions. At this time in vitro or in vivo expression results are available for only a few of the cloned tRNA gene regions.

Delimitation of a transcription unit is typically by a combination of in vivo and in vitro experimentation. The boundaries of a transcription unit are determined in vitro in a relatively straightforward way through manipulation of templates by restriction or deletion and direct characterization of transcripts. On the other hand, identification of tRNA gene start and stop sites utilized in vivo is especially difficult owing to rapid processing of the primary transcript. Because product analysis alone is unreliable, genetic strategies must also be utilized in defining the true limits of a transcription unit. The most direct and reliable information will invariably come from manipulations of the DNA of the transcription unit itself, although a considerable amount of valuable information can and has been developed from mutants defective in RNA processing. However, until mutants exist for all of the processing activities involved, this approach is currently better suited for characterizing processing itself rather than defining primary products. An important exception is the transcriptional start site which is easily recognized because of the unique 5'-triphosphate. Precursors with original 5'-triphosphates often occur in ribonuclease P mutants, but there is not yet any mutant that will yield unprocessed primary transcripts.

Because the bulk of the cloned tRNA genes have become available only recently, in vivo analyses of relevant transcription signals have been carried out in only a few cases. More expression information has been developed from in vitro studies, but even here it is early, with actual termination sites defined for only six of the cloned regions; two of these studies are not yet published. While the in vitro approaches are easier once a sequenced clone is available, these results too must be interpreted with caution. Specific initiation and termination sites can be fairly readily defined in vitro, but it cannot be assumed that these sites necessarily reflect the true in vivo situation. Even when we can expect transcriptional start and stop points identified in vitro to also function in vivo there is the possibility that different or additional signals may be used by the cell. Not only is there the possibility of modulating expression levels for an operon, but even of redefining the transcription unit itself through the use of multiple initiation or termination signals. Suggestions that such might be the case have already surfaced for a few of the genes evaluated. These instances, noted below, serve to underscore the importance of exercising caution in attempting to define transcription units solely from sequencing data, even when coupled with in vitro expression assays. These approaches, while proven generally sound, could in cases yield information that falls well short of reflecting the true in vivo situation. On this cautionary note, we now review the structures of the individual transcription units as defined by the sequence and expression results currently available.

## Organization and Structure of the Individual Transcription Units

Expression analyses have been conducted for 6 of the 15 non-*rrn* tRNA gene clones sequenced. Thus, for these cases it is possible to correlate sequence and transcription results. Our knowledge of the other gene systems is limited to the DNA sequence information and, in a few instances, identification of incompletely processed precursor RNAs. The transcription units best understood at the present are the *metT* (*supB-E*) gene set and the mixed-function operons *tyrT*, *thrU*, and *metY* (*nusA*); preliminary results are available for the *argT* and *serT* gene systems.

The *metT* and *argT* transcription units coordinate the synthesis of seven and four tRNAs, respectively. Both operons are defined at the sequence level by single promoter and rho-independent-like terminator elements that are functional and in good agreement with the canonical consensus sequences (37, 67). No recognizable promoter or terminator sequences occur in the intercistronic regions for either gene cluster. Results from in vitro expression analyses corroborate the structures deduced for each operon from the DNA sequence. Transcripts were observed in each case which corresponded in size to the region defined by the deduced start and stop points (68; L. M. Hsu, personal communication). Evidence that both tDNA clusters function as multigene operons in vivo came from earlier characterization of precursor tRNAs in processing mutants, although in neither case were products corresponding to unprocessed transcripts observed (75, 81, 87).

The situation for the best-characterized mixed-function operons is more complicated. The genomic locus corresponding to *tyrT* consists of a tandem repeat of cistrons encoding $tRNA_1^{Tyr}$. While this locus has received more attention than any other so far, much of this effort has involved characterization of resectioned clones carrying only one tDNA element. Accordingly, there is high-resolution insight into the functioning of portions of this gene set but considerably less for the complete gene cluster. The sequences of the region of the operon corresponding to the promoter, the tDNA itself, and the trailing 3' segment were developed early (3, 22, 84), but it was only recently that complete sequence information became available for the intercistronic spacer and the region upstream of the promoter (62, 77). The sequence and expression results show that the two $tRNA_1^{Tyr}$ genes are served by a common promoter and that the operon also encodes a 33-amino acid, arginine-rich P protein in the region downstream of the second tRNA gene. Full-length transcripts have been observed in both in vivo and in vitro transcription assays (56, 77) and in a coupled cell-free transcription-translation system (3). Interestingly, the in vitro results show the occurrence of a graded series of three well-spaced rho-dependent terminators, one of which separates the second tRNA gene from the P protein-coding region (26, 56). The role of the protamine-like P protein and its significance in a tRNA operon pose particularly fascinating questions for future study.

Substantive genetic and biochemical analyses have also been carried out with the five-gene *thrU* operon which encodes four tRNAs and the translational elongation factor Tu (*tufB*). Here too, coexpression has been demonstrated both in vivo (40, 59) and in vitro (64) with full-length transcripts observed for both situations. In one in vivo study the pattern of expression observed was more complex, however, with evidence of functional minor promoter elements occurring between the second and third tRNA genes and before the *tufB* cistron (59). These results, obtained from gene fusion analyses involving different portions of the operon and the *lacZ* gene or a gene encoding tetracycline resistance, raise the possibility that differential expression of the genes in this operon may occur in vivo.

The sequence and expression results for the *metY* gene region suggest that the cistron coding for $tRNA_{f2}^{Met}$ is also part of a larger mixed-function operon and that this transcription unit too may have the potential for being expressed in differential fashion. The sequence results show the *metY* gene to be followed by an ORF encoding a 15-kilodalton protein and the gene for the *nusA* protein. Two transcripts are produced from the region in vitro in unequal amounts, both apparently from a single promoter immediately upstream from the *metY* gene (51). The major product corresponded to a transcript terminating between *metY* and the ORF element and the second included both the ORF and *nusA* sequences. Only one promoter region can be recognized in the DNA sequence, while familiar rho-independent, terminator-like elements occur in the regions corresponding to both transcriptional endpoints observed. A good Shine-Dalgarno sequence appears before the *nusA* gene, but none is evident upstream for the ORF sequence. Additional evidence suggesting coexpression came from results of maxicell analyses which showed the production of proteins from this region appearing to correspond to both the 15-kilodalton protein and *nusA* (51). Interestingly, subsequent extension of the DNA sequence through the *nusA* gene revealed the neighboring *infB* gene coding for translational initiation factor 2 to be only 24 bp downstream from the *nusA* coding sequence (50). The close proximity of *infB* suggests the possibility that this gene and perhaps even others distal to it may also be part of the *metY* transcription unit.

Less is known about the remaining cloned gene systems. Despite the lack of definitive expression data, however, the sequence results help establish the existence of the aforementioned major classes of tRNA transcription units and provide a preview of yet more interesting and intriguing mechanisms for coordinating tRNA gene expression. As noted above, the cloning and sequence results have indicated new biological roles for tRNA itself, with the *dnaY* and *divE* loci shown to encode arginine and serine tRNAs, respectively (66, 94). The sequence information suggests that the tDNA regions at the *serT* (*divE*), *leuX* ($Su_6^+$), and *proV* loci may also be part of complex operons that encode both tRNA and protein (54a, 94, 108). ORF elements follow the tRNA gene in each case. For the *serT* gene there are cited but unpublished expression data indicating that the 657-bp ORF adjoining the $tRNA_1^{Ser}$ gene directs the synthesis of a 23.5-kilodalton membrane-like protein in vivo (94). No promoter-like sequences can be recognized in the tRNA-ORF spacer regions of any of these latter three clones; however, terminator-like elements of the rho-independent variety are apparent in each case.

Five genomic clones are known thus far to contain only one tRNA coding region, and the sequence results available suggest that three of these may themselves constitute transcription units. The three genes bounded by consensus-like promoter and rho-independent terminator elements are *asnT* (Yoshimura et al., in preparation), *glyW* (S. D. Tucker and E. J. Murgola, Biochimie, in press), and *pheU* (83). The two tDNA segments for which candidate terminators have not yet been identified are the *argU* (*dnaY*) and *serU* loci.

The two remaining cloned tDNAs for which sequence but not expression data are available are the *leuV* and *lysT* genes. Each contains a cluster of three tDNAs (19, 109). The *leuV* locus encodes three identical genes for $tRNA_1^{Leu}$ while the *lysT* region contains a $tRNA^{Val}$ cistron flanked by genes for $tRNA^{Lys}$. Both cloned regions are known to correspond to polycistronic operons based on detection of the related multimeric tRNA precursors in processing mutants (72, 87). The absence of familiar terminator domains in the sequenced DNAs precludes even tentative definition of the boundaries of these operons.

### The tRNA Genes of the *rrn* Operons

The 14 tRNA genes encoded in the seven rRNA operons occur in two regions (Fig. 3). Ten tRNA cistrons have been shown to occur in the 16S to 23S spacer, and the other four tDNAs are in the distal region after the genes for the 23S and 5S RNAs. Two patterns of distribution are observed for the tRNA genes in the spacer region. Three operons, *rrnA*, *rrnD*, and *rrnH*, possess tandem cistrons for $tRNA_1^{Ile}$ and $tRNA_{1B}^{Ala}$ while a single $tRNA_2^{Glu}$ gene occupies the corresponding region of the other four operons (23). tRNA genes are only known to occur in the distal regions of three operons thus far, and the structures of these segments are all different. The distal portion of *rrnC* contains tandem sequences encoding $tRNA_1^{Asp}$ and $tRNA^{Trp}$. The corresponding regions in *rrnD* and *rrnH* each contain one tRNA gene, coding for $tRNA_1^{Asp}$ in both instances. The distal region of the *rrnD* operon is unique among the *rrn* operons thus far in that it contains a second gene for 5S RNA located immediately downstream from the $tRNA_1^{Thr}$ gene (21).

At least partial DNA sequence information has been developed for the tRNA gene regions of five *rrn* operons. Sequences of four spacer tDNA segments are known including both the single- and double-gene arrangements. The sequenced tDNAs and their origins are the $tRNA_2^{Glu}$ gene of the apparently fully sequenced *rrnB* (12), the corresponding segment in *rrnE* (65), and the paired $tRNA_1^{Ile}$-$tRNA_{1B}^{Ala}$ genes of *rrnD* (111, 112). All four tDNAs known to occur in the distal regions have been sequenced (21, 85, 110).

## DOMAINS OF THE tRNA TRANSCRIPTION UNIT

### Promoter Elements

The sequences of the promoter regions of the nonribosomal tRNA transcription units are shown in Fig. 4. Inspection of these regions shows the consensus sequences for the known and likely −35 and −10 elements to agree well with the consensus sequences developed for all *E. coli* genes characterized to date. The consensus sequences for the tRNA gene regions are g TTGAc- for the −35 domain and TAtAaTG for the Pribnow box at −10 from the start of transcription (T, thymine). The corresponding sequences for all genes are tcTTGAca at the −35 element and TAtAaT- for the Pribnow box (34, 76). In the designations adopted here, a large uppercase letter specifies a frequency of about 75% or greater, small uppercase letters correspond to conservation at the 60% level, and lowercase letters represent frequencies of less than 60%.

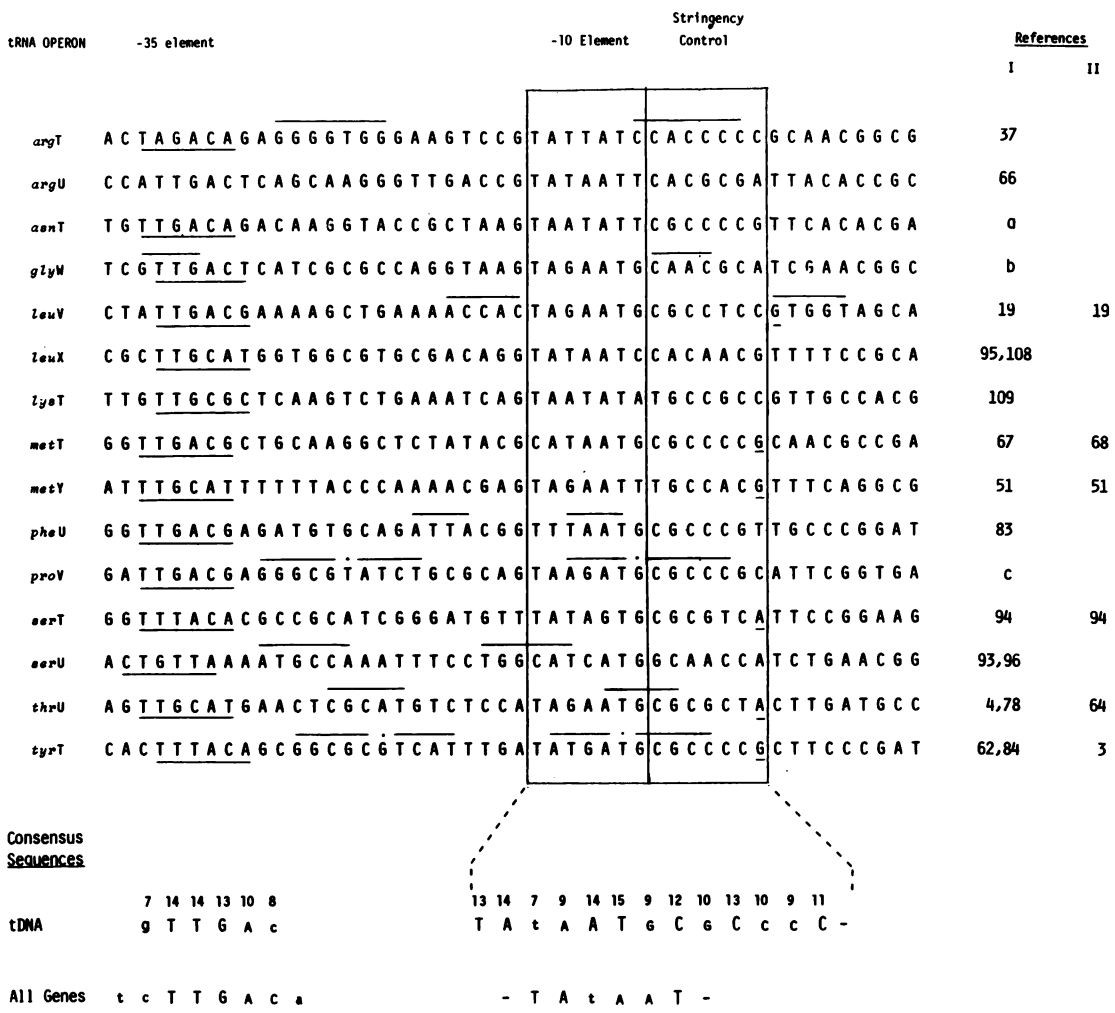



FIG. 4. Sequences of *E. coli* tRNA gene promoter regions. Forty-eight nucleotides of the nontranscribed strand are shown. All sequences are aligned at the −10 domain. The −35 region hexamer is underlined, and the sequences of the −10 domain and proposed stringency control discriminator element are boxed in. Sequence elements with dyad symmetry are identified with overlines. Nucleotides corresponding to demonstrated transcriptional startpoints either in vivo or in vitro are underlined. References in the two groupings correspond to the DNA sequence results (I) and identification of 5′ termini of transcripts (II). Unpublished results were supplied by Yoshimura et al. (*asnT*) (a), Tucker and Murgola (*glyW*) (b), and Kuchino and Nishimura (*proV*) (c). Letter designations in the consensus sequences for the −35, −10, and stringency control regions (34, 98) reflect the different relative frequencies with which a given nucleotide occupies a position; large uppercase letters correspond to a frequency of >75%, small uppercase letters indicate a frequency of >60%, and lowercase letters specify a frequency of >50%. The consensus information for all *E. coli* genes is taken from references 34 and 99.

Not surprisingly, the distance between the −35 and −10 elements is also consistent with the spacing observed for other *E. coli* promoters, ranging from 16 to 18 nucleotides. For all promoters, 90% of the spacings are 17 ± 1 bp (34). In the cases of the 15 tRNA gene promoters, 6 have a spacing of 16 bp, 8 are separated by 17 bp, and one has the two promoter elements positioned 18 nucleotides apart.

A third highly conserved sequence element occurs in the promoter region, corresponding to the suggested discriminator domain common to all *E. coli* genes under stringent control (97, 98). This element consists of a GC-rich 7-bp sequence located immediately 3′ to the −10 region. Transcription of genes subject to stringent control is regulated negatively and in coordinated fashion in response to amino acid limitation (29, 43). Genes known or believed to be under stringent control principally include those encoding the metabolically stable RNAs, several ribosomal proteins, and subunits of RNA polymerase. The consensus sequence for the putative stringency control element in the tRNA gene promoters is CGCccC-, in good correspondence with the conserved sequence described earlier for all promoters subject to the stringent response (97, 98). While the presence of

TABLE 2. Palindromic elements in the promoter region

| tRNA operon | Size of dyad (bp) | Spacing (bp)[a] |
|---|---|---|
| *argT* | 7 | 21 |
| *glyW* | 4 | 30 |
| *leuV* | 5 | 19 |
| *pheU* | 4 | 9 |
| *proV* | 10 | 18 |
| *serU* | 6 | 13 |
| *thrU* | 5 | 17 |
| *tyrT* | 10 | 15 |

[a] The values given are the center-to-center distances between dyad units.

this sequence is well correlated with stringent control, clear-cut direct evidence for its role in the stringent response is still lacking. Perhaps the best evidence for its involvement comes from a recent demonstration showing that deletions in this region in *rrnB* affect inhibition by guanosine tetraphosphate in vitro (54). In the current context, it should be noted that stringent control has been well demonstrated for total tRNA synthesis (43) but has not been well documented for individual tRNA genes to date (98). Thus, the possibility remains that some tRNA genes may not be regulated in this way.

Another particularly interesting feature evident in several of the tRNA gene promoter sequences is the occurrence of inverted repeat elements. The existence of dyad elements in this region suggests a possible functional role in modulating expression. Inspection of the sequences in Fig. 4 show 8 of the 15 promoters to possess dyad elements that could correspond to protein recognition sites. These elements are also identified in Table 2. Inverted repeats of less than four bases per element are not identified here because of a perceived lower potential for being functionally relevant.

In all but one case the downstream element includes nucleotides in the Pribnow box-stringency control region. The one exception is *leuV*, where the rightmost element of the dyad is located immediately 3′ to the stringency control element, corresponding to the first five bases of the transcript. Thus, all of the elements identified are in functionally promising positions. The various promoter dyads contain from 4 to 10 nucleotides per element and are spaced from 9 to 30 bp apart center to center. The interelement spacing in several cases is such that both domains could be on the same face of a normal DNA B helix, consistent with the properties of several DNA-binding proteins (105). There is no evidence implicating a functional role for any of these elements at this time. If a regulatory role should be shown, however, the existence of these elements in some but not all promoters could provide a basis for modulating expression in a differential manner. The knowledge that the synthesis of some (88) but not all tRNAs (McKay and Fournier, manuscript in preparation) is subject to growth-rate-dependent metabolic control raises the possibility that these symmetrical elements could participate in this control process. Consistent with this notion is the knowledge that expression of the *tyrT* and *leuV* genes has been shown to be proportional to growth rate (6, 20) and that both contain inverted repeats in the promoter region.

### Elements Upstream of the Promoter

Results from recent expression studies indicate that elements upstream of the promoter region are also involved in tRNA gene transcription (9, 57, 99). This conclusion is reinforced somewhat by results from sequence analysis that show the presence of conserved sequences, direct repeats, true palindromes, and dyad symmetries, all of which could serve as sites for protein interaction (62, 98). The experimental results that revealed that upstream sequences contribute to transcriptional efficiency are based on both in vivo and in vitro analyses. First evidence came from in vivo expression assays conducted with deletion derivatives of a *tyrT* promoter fragment fused to *galK* (57). Activity assays for a graded series of *Bal* 31 deletions spanning the region −99 to −33 bp from the normal *tyrT* transcriptional start point revealed that sequence elements between −40 and −98 were required for optimal *galK* expression. It was estimated that deletion of this region resulted in a 10-fold reduction in transcriptional activity. Results from subsequent in vitro RNA polymerase binding studies indicated that enzyme could protect segments of this region from deoxyribonuclease I digestion (99). These results, taken with the presence of oppositely oriented −35 and −10 promoter-like elements, were interpreted as evidence for additional polymerase-binding sites, one of moderate strength in the region −44 and −67 and multiple weaker sites further upstream (98, 99).

Another particularly revealing demonstration of the involvement of upstream regions in expression has come from a recent characterization of a tRNA gene set from the genetically similar *S. typhimurium* (9). Although we compare the *E. coli* and *Salmonella* tRNA genes later in this review, this particular study is included at this point since it reflects the presence of an important transcriptional element that could also exist in *E. coli*. The results, obtained with a tRNA operon that has a near-identical homolog in *E. coli* confirm the occurrence of a positive effector region upstream from the promoter and identify a DNA conformational effect that is associated with this region. In this study, comparative structure-expression analyses were carried out with cloned genomic DNA segments containing wild-type and mutant alleles of the *hisR* locus. The *hisR* mutation was shown previously to be correlated with reduced levels of $tRNA^{His}$ in vivo (10). This region is now known to encode a tRNA operon with tandem genes for $tRNA^{Arg}$, $tRNA^{His}$, $tRNA_1^{Leu}$, and $tRNA^{Pro}$ (8); the related homolog in *E. coli* is the gene set designated *argT* (37).

Comparisons of the DNA of wild type and two different *hisR* mutants revealed the mutations to correspond to a three-bp deletion at a position −70 from the transcriptional start site in one case and an alteration in the −10 region for the second (9). The occurrence of the −70 deletion results in an approximate two-fold reduction of $tRNA^{His}$ synthesis both in vivo and in vitro. Interestingly, the electrophoretic mobilities of restriction fragments encompassing the −70 region were observed to be very different for wild-type and mutant DNAs. In these analyses the wild-type DNA exhibited abnormally low mobility in nondenaturing conditions, similar in effect to the behavior of DNA believed to be bent (33, 107). The mutant DNA, on the other hand, did not show this unusual property. These results are consistent with those observed in the aforementioned *tyrT* deletion analysis and, together, identify another DNA feature of promising importance.

Inspection of sequences upstream from the primary promoter of several *E. coli* genes, including tRNA and non-tRNA cistrons, has revealed the occasional presence of additional promoter-like elements as noted above for *tyrT* and regions of apparently conserved AT nucleotides (98). The −35 and −10 promoter-like elements occur alone and together, on both strands in both orientations, and with variable spacing. Despite this complex distribution, the presence of these elements has served as a basis for speculation about possible secondary sites for polymerase binding (57, 100). Whether or not such domains have any role in tRNA gene expression must, of course, await the results of further experimentation. Blocks of AT nucleotides have also been noted in regions upstream from the promoters of some tRNA genes and also promoters of the *rrn* operons (98). As with the still hypothetical secondary promoter elements, knowledge about the significance of this particular recurring feature must also await experimental analysis.

TABLE 3. Sequences of known tRNA gene terminator regions[a]

| Operon | Distal gene | Space (bp) | Termination region | Reference |
|---|---|---|---|---|
| *argT* | $tRNA^{Pro}$ | 0 | -ATTTTGAACCCCGCTTCGGCGGGGTTTTTTG- | [b] |
| *metT* | $tRNA_2^{Gln}$ | 0 | -CATTAAAAAAGCTCGCTTCGGCGAGCTTTTTGCTTTTC- | 68 |
| *metY* | $tRNA_{f2}^{Met}$ | 119 | -TAAAGCCCCGATTTATCGGGGTTTTTTGTTATCTGACTACAGAATAACT GGGCTTTAGGCCCTTTTTTTA- | 51 |
| *serT* | $tRNA_1^{Ser}$ | 3 | -AAGATAAGGGGTTAGCTAAATGCTAACCCCTTTTTCTTTTGCCTGTC GAAATTCTCAGGGCGTTATATTTG- | [c] |
| *thrU* | *tufB* | 25 | -GCACTAAAGGGCATCATTGATGCCCTTTTTGCACG- | 40, 64 |
| *tyrT* | $tRNA_1^{Tyr}$ | 210 | -ATCCGCAATCAAATATTCTGCCCAT- | 22, 56, 77 |

[a] The sequences of the regions in which transcriptional termination has been shown to occur are presented. Listed are: the last gene preceding the termination site, the distance between this gene and the terminator region, and the sequence of the region encompassing the termination point. Inverted sequences are overlined, and the actual site where termination is known to occur (*) is shown for two cases. All sites shown are utilized in vitro; only the *thrU* and *tyrT* sites are known to also function in vivo.

[b] L. Hsu, unpublished data.

[c] Unpublished data described in reference 94.

### Transcribed Spacer Regions

All *E. coli* tRNAs are believed to be derived by processing of larger precursor molecules containing stretches of additional nucleotides at both ends of the mature tRNA sequence (73, 87). In the cases of the complex operons that also encode a protein or rRNA, the primary transcript obviously includes these products as well. Ignoring the non-tRNA products, the number of nucleotides in the spacers flanking the individual tRNAs is variable. In transcription units where the first gene is for a tRNA, the 5′ spacers range in size from about 4 to 66 bases (Fig. 2). The sizes of the intercistronic spacers are equally diverse, ranging from only 2 to more than 200 nucleotides in length. The lengths of the 3′ terminal spacers are less well defined because of uncertainty about the actual sites where transcriptional termination occurs. As will be detailed below, termination sites have been identified in only a few cases and with variable precision. For the genes in this subset the number of nucleotides separating the termination signal from the 3′ ends of the upstream tRNA is estimated to vary from about 30 to more than 200 nucleotides. A similar range holds for the 3′ spacers of several of the less-characterized genes in which candidate rho-independent terminator elements have been identified in the DNA sequence (Fig. 2).

Aside from being substrates for processing ribonucleases, no other function has yet been ascribed to any of the intercistronic spacer elements. Inspection of these sequences reveals several to possess inverted repeat elements and conserved sequences, and some are enriched in AT bases. While processing is known to occur at some stem-loop elements (40), neither this nor any other feature appears to be constant to all tDNA spacers. The apparent absence of strictly conserved features is consistent with the knowledge that several enzymes are involved in processing. Thus, processing with certain spacers may involve recognition of stem-loop domains and conserved sequence features, but perhaps relies heavily or even exclusively on the occurrence of the adjoining tRNA in others. It seems almost inescapable that the tRNA moiety itself is involved in the extreme case of a spacer in the *lysT* transcript that consists of only two bases.

### Terminator Elements

Transcription termination for the tRNA genes of *E. coli* appears to occur by both rho-dependent and factor-independent mechanisms, and perhaps by attenuation. DNA regions where termination is known to occur have been identified by in vivo or in vitro methods for six operons, only four of which have been reported thus far. The sequences of these regions are provided in Table 3. In five of the six cases termination occurs at an inverted repeat that closely resembles the generalized rho-independent terminator element (36, 76). The exception is the *tyrT* operon, where primary termination is apparently rho dependent, occurring at AT-rich elements present in all copies of a 178-bp sequence that is tandemly repeated 3.14 times (22, 56).

While termination sites have not yet been identified for the remaining operons, the DNA sequence results indicate that several will likely be of the rho-independent variety. In addition to the five terminators of this class identified by direct product analysis, strong candidates can be found in the DNA sequences of five other transcription units. The locations of these hypothetical terminators are shown in Fig. 2.

In addition to the terminators that define the limits of the various transcription units, it is also possible that some operons contain attenuator elements. Current candidates for possessing this control feature include some of the complex operons that encode both tRNA and protein; of course, others could also benefit from this means of modulating expression. While attentuation is not known to be involved in the expression of any tRNA operon at this time, the early expression and sequence results obtained for four gene systems do provide a basis for suggesting this possibility. The relevant operons are *metY* (*nusA*), *serT*, *leuX*, and *proV*. Inasmuch as *nusA* is known to enhance termination at rho-independent sites (36, 105), the *metY* operon encoding *nusA* and probably *infB* is the strongest candidate. As noted above, two transcripts are produced in vitro from this region, both from the same promoter. Most of the RNA derived is from the *metY* gene itself, with the remainder corresponding to joint transcripts terminating after *nusA*. The potential for obtaining even longer transcripts terminat-

ing after the adjacent *infB* gene has not been evaluated. While there is real risk of overextending such in vitro results, the occurrence of both mono- and polycistronic transcripts from the same promoter and the presence of recognizable terminator-like elements after each gene is at least consistent with the possibility that this gene system could be subject to attentuation control in vivo.

The transcription units encoding *serT*, *leuX*, and *proV* are also possible candidates for attenuation. In each case the single tRNA gene is followed by a recognizable terminator-like sequence and downstream ORF or protein gene. In the case of *serT*, transcription termination occurs after the tRNA gene in vitro, and no promoter-like elements are evident in the spacer region separating the two coding elements. If the known or potential protein cistrons associated with these tRNA genes are shown to be part of the same transcription unit, then the necessary structural features for attenuation are in place. No evidence exists for differential expression through antitermination. Nonetheless, the organization and early expression results suggest that this possibility should be given consideration.

### Do Elements Distal to the Terminator Have a Role in Expression?

An intriguing observation exists in the literature that suggests the possibility that information distal to a tRNA gene might also influence expressional activity (1, 80). A comparison of in vivo levels of tRNA in cells containing cloned synthetic and natural *tyrT* genes showed that cells harboring the synthetic gene contained markedly more $tRNA^{Tyr}$ than those with the natural gene. In both cases the analysis involved expression of plasmid-encoded tDNA. While the basis for the reduced synthesis from the natural gene was not determined, a potentially important factor was a difference in the amount of DNA that naturally flanks the *tyrT* gene. The synthetic gene was about 200 nucleotides in length, including 56 bp of the promoter region and some 25 bp beyond the bases corresponding to the 3′ end of the mature tRNA. The natural gene, on the other hand, was in the central portion of a 1,200-bp fragment of genomic DNA. Thus, normal chromosomal sequences flanked the natural gene, extending well beyond the upstream region known to enhance expression and continuing some 560 bp beyond the end of the encoded tRNA. While other factors may be responsible for the different amount of tRNA observed, it has been reasonably suggested that the effect could be due to the presence or absence of the normal flanking sequences (1). If elements upstream of this promoter serve only to enhance expression, then the possibility exists that elements in the downstream region could be responsible for the reduced expression levels observed with the natural gene. This distal region includes both a 178-bp repeating sequence that begins within the tDNA and the adjacent coexpressed ORF which encodes the small P protein.

## COMPARISON OF THE tRNA GENES OF *ESCHERICHIA COLI*, *SALMONELLA TYPHIMURIUM*, and *BACILLUS SUBTILIS*

Until recently little detailed information has been available for the tRNA genes of procaryotes other than *E. coli*. Happily, that situation no longer exists. Especially good progress has been made in the past few years in characterizing the genes from *B. subtilis*. In addition, early reports about the structure of a tRNA operon from *S. typhimurium* confirm that the strong genetic similarity this organism shares with *E. coli* extends to the tRNA genes as well. In this section we compare the major features of the tRNA genes of *E. coli* with those of *S. typhimurium* and *B. subtilis*. Our comparative evaluation is restricted to these procaryotes alone as too little information is available from other sources to permit generalizations to be made.

### The tRNA genes of *S. typhimurium*

After *E. coli*, the gram-negative organism about which most is known genetically and biochemically is *S. typhimurium*. Considerable evidence is available indicating that these two organisms are closely related genetically. This includes strong similarity of their genetic maps, substantial relatedness of the genomic DNAs in cross-hybridization analyses, and detailed comparisons of several physiological processes (82). In the current context the genetics and biochemistry of the tRNAs and aminoacyl-tRNA synthetases also show a high degree of relatedness. Of six *S. typhimurium* tRNAs sequenced thus far, four are known to correspond precisely to species present in *E. coli* (90). As noted earlier, characterization of the DNA of the *hisR* locus of *S. typhimurium* has revealed it to be almost exactly like the four-gene *argT* operon of *E. coli* (8, 37). Over 95% sequence homology exists for a 600-bp region that starts 100 nucleotides upstream of the promoter and extends some 50 nucleotides beyond the likely terminator. Three of the four tDNAs are identical, and the fourth, $tRNA^{Arg}_{CCG}$, differs by only two nucleotides. The −35 domain of the promoters differ by one base, while the −10 domains and the positive stringency control regions and terminator elements are identical for the two operons. Both gene sets map in the same region of the respective chromosome with the *Salmonella hisR* locus at 82 min on the normalized 100-min map and the *argT* cluster in the region of 82 to 89 min on the *E. coli* map (44, 82). The essentially identical nature of these two operons suggests that the documented genetic similarity between these organisms can be expected to extend to many, perhaps even most, of the tRNA genes.

In the context of the current review, a point of major significance mentioned earlier is the possibility that the upstream control elements recently discovered for the *Salmonella hisR* locus may also be common to the tRNA genes of *E. coli*. Interestingly, the sequences of the *Salmonella* and *E. coli* operons differ at the −70 site where the upstream *hisR* mutation was shown to occur (9). This mutation corresponds to a deletion of three bp, only one base of which is common to the *E. coli* operon. If both operons possess the upstream transcriptional element, the difference in sequence suggests that the structural features that define this element may not be limited to the small segment altered by the *hisR* mutation. Thus, the three-base-pair deletion may affect other features that may be in common between the two organisms, such as a higher-order structure or spacing between sequence elements that are conserved. The aforementioned dramatic difference in electrophoretic mobility observed for wild-type and mutant restriction fragments encompassing this region demonstrates that this site does indeed influence higher-order structure.

### The tRNA genes of *B. subtilis*

The gram-positive bacterium *B. subtilis* is the procaryotic organism after *E. coli* about which most information on tRNA genes is available. A review on the structure and organization of the tRNA genes in this organism appeared in

a recent issue of this publication (103). In this section the major features of the tRNA genes of *B. subtilis* are compared with those of *E. coli*. All references to the *B. subtilis* work are found in the review by Vold (103).

The number of tRNA genes in *B. subtilis* has been estimated at about 60, in general agreement with the estimates for *E. coli*. A total of 51 genes from *B. subtilis* have been sequenced, and the occurrence of another 10 has been estimated by hybridization. Approximately 40 tRNA species have been detected by chromatographic and two-dimensional electrophoretic procedures. This number compares with 50 to 60 species observed in two-dimensional gel assays of *E. coli* tRNA (McKay and Fournier, unpublished data). There is uncertainty about the estimates for both organisms owing to incomplete resolution and the possibility that some resolved species may differ only at the level of base modification.

The manner in which the tRNA genes are organized in the respective chromosomes differs considerably for the two organisms. The genomic clustering observed for *E. coli* is even more pronounced in *B. subtilis*. Six gene sets have been characterized by sequence analysis in *B. subtilis*, and these are clustered mainly in two regions of the chromosome. Of these only one set appears not to be associated with rRNA genes. Of a total of 10 *rrn* gene sets, 3 have been analyzed. Two strongly resemble the *rrn* operons of *E. coli* with genes for $tRNA^{Ile}$ and $tRNA^{Ala}$ in the 16S to 23S spacer region. Two of the tRNA gene sets are much larger than any known for *E. coli*, with 16 and 21 tandemly linked cistrons.

Comparison of the structures of the tRNA gene sets for the two organisms reveals interesting differences. At the organizational level there is no case yet in *B. subtilis* in which a tRNA gene is repeated in a gene cluster, a situation that occurs for several operons in *E. coli*. Examination of the various coding regions in *B. subtilis* shows the 3'-terminal -CCA sequence common to all tRNA to be absent in 12 of the 51 genes sequenced. This deficiency occurs in every cluster evaluated thus far except those in the spacer region of the *rrn* operons. This situation occurs for certain of the tRNA genes encoded by some of the coliphage but is not a property of the *E. coli* genes themselves (27).

When all genes for *B. subtilis* and *E. coli* are considered, the promoter sequences in *B. subtilis* are seen to be more diverse than those of *E. coli*. This complexity, which is believed to be related to the multiplicity of RNA polymerases in this organism, has not yet been seen in the putative promoters for the tRNA genes. Here, only one class of promoter is known thus far, and it appears to resemble the consensus sequences developed for the *E. coli* promoters. Interestingly, both promoter- and attenuator-like sequences are found within gene sets of *B. subtilis*. Similarly, a region analogous to the proposed stringency control domain in *E. coli* is evident for some of the presumed *Bacillus* promoters, including one that is internal in the unprecedented 21-gene cluster. A notable difference from the *E. coli* element, however, is that the first base of the candidate stringency control region corresponds to the startpoint of transcription.

As with the *E. coli* gene sets, the number of nucleotides in the intercistronic spacer region varies broadly in *B. subtilis*, ranging from 2 to 81 bp. Also like *E. coli*, the presence of a spacer as small as two bp argues that processing in this organism, too, must in cases rely on the presence of the tRNA structure itself.

Candidate terminator elements have been identified in *B. subtilis* that are similar to the known and hypothetical factor-independent terminators of *E. coli*. Using the same criteria, a potential attenuator element has also been identified in *B. subtilis* between the second and third cistrons of a six-gene set. Expression analyses with the assorted cloned *B. subtilis* genes now available hold promise for revealing several interesting new features of tRNA biosynthesis, some possibly related to the complex sporulation process.

## DISCUSSION

In this section we comment on future approaches to studying *E. coli* tRNA gene structure and function and identify major outstanding issues yet to be resolved.

### Future Approaches to Cloning and Mapping

The various strategies that have been used thus far for detecting and mapping tRNA genes will no doubt be applied to future efforts in this direction as well. Detection of tDNA loci through complementation of tRNA functions could well reveal even more new biological roles for tRNA. However, the use of individual clones to relieve blocks in more traditional activities such as decoding or synthetase function will likely be more restricted than previously and require the use of more sophisticated genetic approaches. Strategies for identifying new suppressors of nonsense mutations will be more complicated as the bulk of the easily developed *sup* genes with this function have presumably already been identified, with the possible exception of some UGA (opal) suppressors.

Suppression of thermolabile synthetases with cloned tDNA can be expected to work for some but not all classes of ts mutants. A cognate tRNA may not be able to provide protection in certain cases because of the nature of the protein defect. In other cases protection may require strong overproduction of the tRNA, possibly necessitating the use of expression vectors. This approach is also predictably more complicated when multiple isoacceptor species and repeated genes enter into consideration. Even if all genes for the desired acceptor activity have been detected there is the matter of distinguishing among the various complementing tDNAs. One solution to the discrimination problem could be the use of specific hybridization probes to identify the tDNA clone of interest. These assays would, of course, require that pure product tRNA or a corresponding DNA or RNA oligonucleotide probe be available.

An approach to tDNA cloning that has not yet been used but could be useful for tRNAs of low abundance is the use of an in vivo expression assay that is sensitive to enhanced production of the tRNA at issue. Here, a gene encoding a messenger RNA with the corresponding rare codon would be introduced into the host cell used in screening. Ideally, expression would be low in the absence of the cloned tDNA but detectably improved in its presence. The success of this approach will depend on two vital conditions, limiting tRNA in the host cell and observable phenotypic effects on overproduction of the cloned tDNA. In this last regard, discrimination in the screening assay may well require an amplified physiological response to small changes in tRNA levels. The first condition may be met if the tRNAs for the codons known to be rare in *E. coli* (32, 41, 45) are correspondingly low, especially if the test messenger RNA is highly abundant. A low base-line activity will be essential as the extent of tRNA overproduction by the clone may only be modest. In one previous case production of the tRNA was enhanced by only two to threefold in transformants containing tDNA amplified some 10- to 20-fold (37).

In cases in which a suitable hybridization probe is available, the task of identifying desired tRNA gene clones can be made markedly easier if the tDNA is enriched before cloning. This approach has been used by one of our laboratories to clone the *E. coli* 4.5S and 6S RNA genes (38, 39). In these cases fragments carrying the DNA of interest were enriched 60- to 80-fold by two successive cycles of restriction and electrophoretic fractionation. With this strategy positive transformants were found in the colony banks at the statistically expected level of 1 per 25 to 30 candidates.

The map position of cloned tDNA genes can now be determined with relative ease by using a marker gene placed adjacent to the tDNA sequences in the chromosome. This aim is met by incorporating the marker gene, say, an antibiotic resistance determinant, into the segment of cloned bacterial DNA in vitro in a nonessential region and then effecting an in vivo exchange of this region for the corresponding chromosomal sequence via homologous recombination. Once the substitution strain has been constructed mapping can be done by conventional methods. While this strategy has not yet been applied to mapping of tRNA genes, it has been used successfully for locating the genes for the 4.5S and 6S RNAs (13, 58).

### Analysis of tRNA Gene Expression

The combined use of in vitro methods for direct manipulation of cloned tRNA transcription units and in vivo and in vitro expression analyses provides the most effective means of extending our knowledge about tRNA synthesis (and turnover) in *E. coli*. Important aims for the future include full definition of the individual genes and operons and characterization of the mechanisms by which transcription is regulated. Progress in characterizing the stringent control process has been disappointingly slow, especially in light of the ready availability of cloned operons subject to this control. Indeed, only a little experimental evidence exists for the involvement of the still largely hypothetical discriminator domain (54), despite knowledge about its commonality for several years (97).

Metabolic or growth-rate-dependent control (69) will surely be the subject of considerable attention in the future. Approaches for studying this and other control phenomena will no doubt be based on genetic strategies. A common scheme will involve in vitro fusion of the control region to well-characterized gene systems such as *lacZ* and *galK* (6, 20) coupled with both in vivo and in vitro mutagenesis strategies.

Characterization of the nucleolytic processing steps involved in the conversion of tRNA gene transcripts to mature tRNA is another area in which major advances can be anticipated from the application of molecular genetic strategies. Overall progress in this area has been slow for lack of suitable substrates for identifying and characterizing the specific nucleases involved. With the recent development of in vitro transcription vectors utilizing the promoters of *Salmonella* phage SP6 (14, 53, 63) and phage T7 (61), it should now be possible to generate microgram quantities of specific precursor tRNAs with relative ease. On another tack, it will be of obvious benefit to processing studies to develop strains which overproduce individual processing activities for biochemical characterization. Of course, expression of the corresponding genes may have to be kept in careful check during culturing and only induced just before harvesting of the cells. Cloning of the genes for at least some processing enzymes may have to await identification of the activity through in vitro biochemical analysis and use of immunological rather than functional screening methods. Alternatively, it seems likely that suitable tRNA substrates can be engineered to permit the use of in vivo selection methods in cloning.

Future studies of the *E. coli* tRNA genes seem assured of being especially enlightening and exciting. Clones are now available for a considerable variety of functionally and structurally unique gene systems, including tRNAs with different biological roles and genes subject to different control mechanisms. With the powerful experimental tools already available, it seems all but certain that the near future will bring major new insights to our understanding of the structure and function of these most fascinating genes.

### ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM 19351 from the National Institutes of Health (M.J.F.) and a Grant-in-Aid for Special Project Research from the Ministry of Education, Science and Culture of Japan (H.O.).

We thank our colleagues who shared unpublished information, B. Bachmann for help in developing the refined genetic nomenclature proposed, and C. Hayward for preparing the manuscript and compiling the tRNA gene sequences.

### LITERATURE CITED

1. **Abelson, J.** 1980. The organization of tRNA genes, p. 211–219. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
2. **Altman, S., P. Model, G. H. Dixon, and M. A. Wosniak.** 1981. An *E. coli* gene coding for a protamine-like protein. Cell **26**:299–304.
3. **Altman, S., and J. D. Smith.** 1971. Tyrosine tRNA precursor molecule polynucleotide sequence. Nature (London) New Biol. **233**:35–39.
4. **An, G., and J. D. Friesen.** 1980. The nucleotide sequence of *tufB* and four nearby tRNA structural genes of *Escherichia coli*. Gene **12**:33–39.
5. **Bachmann, B. J.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47**:180–230.
6. **Berman, M. L., and J. Beckwith.** 1979. Fusions of the *lac* operon to the transfer RNA gene *tyrT* of *Escherichia coli*. J. Mol. Biol. **130**:285–301.
7. **Björk, G. R.** 1984. Modified nucleosides in RNA—their formation and function, p. 291–330. *In* D. Apirion (ed.), Processing of RNA. CRC Press, Inc., Boca Raton, Fla.
8. **Bossi, L.** 1983. The *hisR* locus of *Salmonella*: nucleotide sequence and expression. Mol. Gen. Genet. **192**:163–170.
9. **Bossi, L., and D. M. Smith.** 1984. Conformational change in the DNA associated with an unusual promoter mutation in a tRNA operon of *Salmonella*. Cell **39**:643–652.
10. **Brenner, M., and B. N. Ames.** 1972. Histidine regulation in *Salmonella typhimurium*. IX. Histidine transfer ribonucleic acid of the regulatory mutants. J. Biol. Chem. **247**:1080–1088.
11. **Brenner, D. J., M. J. Fournier, and B. P. Doctor.** 1970. Isolation and partial characterization of the transfer ribonucleic acid cistrons from *Escherichia coli*. Nature (London) **227**:448–451.
12. **Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller.** 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. J. Mol. Biol. **148**:107–127.
13. **Brown, S., and M. J. Fournier.** 1984. The 4.5S RNA gene of *Escherichia coli* is essential for cell growth. J. Mol. Biol. **178**:533–550.
14. **Butler, E. T., and M. J. Chamberlin.** 1982. Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. J. Biol. Chem. **257**:5772–5778.
15. **Caillet, J., J. A. Plumbridge, M. Springer, J. Vacher, C.**

**Delamerche, R. H. Buckingham, and M. Grunberg-Manago.** 1983. Identification of clones carrying an *E. coli* $tRNA^{Phe}$ gene by suppression of phenylalanyl-tRNA synthetase thermosensitive mutants. Nucleic Acids Res. **11**:727–736.

16. **Campen, R. K., G. L. Duester, W. M. Holmes, and J. M. Young.** 1980. Organization of transfer ribonucleic acid genes in the *Escherichia coli* chromosome. J. Bacteriol. **144**:1083–1093.
17. **Cortese, R.** 1979. The role of tRNA in regulation, p. 401–432. *In* R. F. Goldberger (ed.), Biological regulation and development, vol. I: Gene expression. Plenum Publishing Corp., New York.
18. **Demerec, M., E. A. Adelberg, A. J. Clark, and P. E. Hartman.** 1966. A proposal for a uniform nomenclature in bacterial genetics. Genetics **54**:61–76.
19. **Duester, G., R. K. Campen, and W. M. Holmes.** 1981. Nucleotide sequence of an *Escherichia coli* $tRNA_1^{Leu}$ operon and identification of the transcription promoter signal. Nucleic Acids Res. **9**:2121–2139.
20. **Duester, G., R. M. Elford, and W. M. Holmes.** 1982. Fusion of the *Escherichia coli* $tRNA_1^{Leu}$ promoter to the *galK* gene: analysis of sequences necessary for growth-rate dependent regulation. Cell **30**:855–864.
21. **Duester, G. L., and W. M. Holmes.** 1980. The distal end of the ribosomal RNA operon *rrnD* of *Escherichia coli* contains a $tRNA_1^{Thr}$ gene, two 5S rRNA genes and a transcription terminator. Nucleic Acids Res. **8**:3793–3807.
22. **Egan, J., and A. Landy.** 1978. Structural analysis of the $tRNA_1^{Tyr}$ gene of *Escherichia coli*: a 178 base pair sequence that is repeated 3.14 times. J. Biol. Chem. **253**:3607–3622.
23. **Ellwood, M., and M. Nomura.** 1982. Chromosomal locations of the genes for rRNA in *Escherichia coli* K-12. J. Bacteriol. **149**:458–468.
24. **Fleck, E. W., and J. Carbon.** 1975. Multiple gene loci for a single species of glycine transfer ribonucleic acid. J. Bacteriol. **122**:492–501.
25. **Fournier, M. J., W. L. Miller, and B. P. Doctor.** 1974. Clustering of tRNA cistrons in *Escherichia coli* DNA. Biochem. Biophys. Res. Commun. **60**:1148–1154.
26. **Fournier, M. J., E. Webb, and S. Tang.** 1977. *In vitro* biosynthesis of functional *Escherichia coli* $Su_3^+$ tyrosine transfer RNA. Biochemistry **16**:3608–3618.
27. **Fukada, K., and J. Abelson.** 1980. DNA sequence of a T4 phage transfer RNA gene cluster. J. Mol. Biol. **139**:377–391.
28. **Gallagher, P. J., I. Schwartz, and D. Elseviers.** 1984. Genetic mapping of *pheU*, an *Escherichia coli* gene for phenylalanine tRNA. J. Bacteriol. **158**:762–763.
29. **Gallant, J. A.** 1979. Stringent control in *E. coli*. Annu. Rev. Genet. **13**:393–415.
30. **Gegenheimer, P., and D. Apirion.** 1981. Processing of procaryotic ribonucleic acid. Microbiol. Rev. **45**:502–541.
31. **Ghysen, A., and J. E. Celis.** 1974. Joint transcription of two $tRNA_1^{Tyr}$ genes from *Escherichia coli*. Nature (London) **249**:418–421.
32. **Gouy, M., and C. Gautier.** 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. **10**:7055–7074.
33. **Hagerman, P. J.** 1984. Evidence for the existence of stable curvature of DNA in solution. Proc. Natl. Acad. Sci. USA **81**:4632–4636.
34. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
35. **Hill, C. W., C. Squires, and J. Carbon.** 1970. Glycine transfer RNA of *Escherichia coli*. I. Structural genes for two glycine tRNA species. J. Mol. Biol. **52**:557–569.
36. **Holmes, W. M., T. Platt, and M. Rosenberg.** 1983. Termination of transcription in *E. coli*. Cell **32**:1029–1032.
37. **Hsu, L. M., H. J. Klee, J. Zagorski, and M. J. Fournier.** 1984. Structure of an *Escherichia coli* tRNA operon containing linked genes for arginine, histidine, leucine, and proline tRNAs. J. Bacteriol. **158**:934–942.
38. **Hsu, L. M., J. Zagorski, and M. J. Fournier.** 1984. Cloning and sequence analysis of the *Escherichia coli* 4.5S RNA gene. J. Mol. Biol. **178**:509–531.
39. **Hsu, L. M., J. Zagorski, Z. Wang, and M. J. Fournier.** 1985. The *Escherichia coli* 6S RNA gene is part of a dual-function transcription unit. J. Bacteriol. **161**:1162–1170.
40. **Hudson, L., J. Rossi, and A. Landy.** 1981. Dual function transcripts specifying tRNA and mRNA. Nature (London) **294**:422–427.
41. **Ikemura, T.** 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. **146**:1–21.
42. **Ikemura, T.** 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for synonymous codon choice that is optimal for the *E. coli* translational system. J. Mol. Biol. **151**:389–409.
43. **Ikemura, T., and J. E. Dahlberg.** 1973. Small ribonucleic acids of *Escherichia coli*. II. Noncoordinate accumulation during stringent control. J. Biol. Chem. **248**:5033–5041.
44. **Ikemura, T., and H. Ozeki.** 1977. Gross map location of *Escherichia coli* transfer RNA genes. J. Mol. Biol. **117**:419–446.
45. **Ikemura, T., and H. Ozeki.** 1983. Codon usage and transfer RNA contents: organism-specific codon-choice patterns in reference to the isoacceptor contents. Cold Spring Harbor Symp. Quant. Biol. **47**:1087–1097.
46. **Ikemura, T., Y. Shimura, H. Sakano, and H. Ozeki.** 1975. Precursor molecules of *Escherichia coli* transfer RNAs accumulated in a temperature-sensitive mutant. J. Mol. Biol. **96**:69–86.
47. **Ilgen, C., L. L. Kirk, and J. Carbon.** 1976. Isolation and characterization of large transfer ribonucleic acid precursors from *Escherichia coli*. J. Biol. Chem. **251**:922–929.
48. **Inokuchi, H., M. Kodaira, F. Yamao, and H. Ozeki.** 1979. Identification of transfer RNA suppressors in *Escherichia coli*. II. Duplicated genes for $tRNA_2^{Gln}$. J. Mol. Biol. **132**:663–677.
49. **Inokuchi, H., F. Yamao, H. Sakano, and H. Ozeki.** 1979. Identification of transfer RNA suppressors in *Escherichia coli*. I. Amber suppressor $Su_2^+$, an anticodon mutant of $tRNA_2^{Gln}$. J. Mol. Biol. **132**:649–662.
50. **Ishii, S., M. Ihara, T. Maekawa, Y. Nakamura, H. Uchida, and F. Imamoto.** 1984. The nucleotide sequence of the cloned *nusA* gene and its flanking region of *Escherichia coli*. Nucleic Acids Res. **12**:3333–3342.
51. **Ishii, S., K. Kuroki, and F. Imamoto.** 1984. $tRNA_{f2}^{Met}$ gene in the leader region of the *nusA* operon in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **81**:409–413.
52. **Jinks-Robertson, S., R. L. Gourse, and M. Nomura.** 1983. Expression of rRNA and tRNA genes in *Escherichia coli*: evidence for feedback regulation by products of rRNA operons. Cell **33**:865–876.
53. **Kassavetis, G. A., E. T. Butler, D. Roulland, and M. J. Chamberlin.** 1982. Bacteriophage SP6-specific RNA polymerase. II. Mapping of SP6 DNA and selective *in vitro* transcription. J. Biol. Chem. **257**:5779–5788.
54. **Kingston, R. E.** 1983. Effects of deletions near *Escherichia coli rrnB* promoter $P_2$ on inhibition of *in vitro* transcription by guanosine tetraphosphate. Biochemistry **22**:5249–5254.

54a. **Kuchino, Y., F. Mori, and S. Nishimura.** 1985. Structure and transcription of the $tRNA1_1^{Pro}$ gene from *Escherichia coli*. Nucleic Acids Res. **13**:3213–3220.

55. **Küpper, H., R. Contreras, A. Landy, and H. G. Khorana.** 1975. Promoter-dependent transcription of $tRNA_1^{Tyr}$ genes using DNA fragments produced by restriction enzymes. Proc. Natl. Acad. Sci. USA **72**:4754–4758.
56. **Küpper, H., T. Sekiya, M. Rosenberg, J. Egan, and A. Landy.** 1978. A rho-dependent termination site in the gene coding for tyrosine tRNA $Su_3$ of *Escherichia coli*. Nature (London) **272**:423–428.
57. **Lamond, A. I., and A. A. Travers.** 1983. Requirement for an upstream element for optimal transcription of a bacterial tRNA gene. Nature (London) **305**:248–250.
58. **Lee, C. A., M. J. Fournier, and J. Beckwith.** 1985. The 6S RNA of *Escherichia coli* 6S RNA is not essential for growth or

protein secretion. J. Bacteriol. **161**:1156–1161.
59. **Lee, J. S., G. An, and J. D. Friesen.** 1981. Location of the *tufB* promoter of *E. coli*: cotranscription of *tufB* with four transfer RNA genes. Cell **25**:251–258.
60. **Lund, E., J. E. Dahlberg, and C. Guthrie.** 1980. Processing of spacer tRNAs from ribosomal RNA transcripts of *Escherichia coli*, p. 123–137. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
61. **McAllister, W. T., C. Morris, A. H. Rosenberg, and F. W. Studier.** 1981. Utilization of bacteriophage T7 late promoters in recombinant plasmids during infection. J. Mol. Biol. **153**:527–544.
62. **McCorkle, G. M., and S. Altman.** 1982. Large deletion mutants of *Escherichia coli* $tRNA_1^{Tyr}$. J. Mol. Biol. **155**:83–103.
63. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
64. **Miyajima, A., M. Shibuya, Y. Kuchino, and Y. Kaziro.** 1981. Transcription of the *E. coli tufB* gene: cotranscription with four tRNA genes and inhibition by guanosine-5′-diphosphate-3′-diphosphate. Mol. Gen. Genet. **183**:13–19.
65. **Morgan, E. A., T. Ikemura, L. E. Post, and M. Nomura.** 1980. tRNA genes in rRNA operons of *Escherichia coli*, p. 259–266. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
66. **Mullin, D. A., G. M. Garcia, and J. R. Walker.** 1984. An *E. coli* DNA fragment 118 base pairs in length provides $dnaY^+$ complementing activity. Cell **37**:669–674.
67. **Nakajima, N., H. Ozeki, and Y. Shimura.** 1981. Organization and structure of an *E. coli* tRNA operon containing seven tRNA genes. Cell **23**:239–249.
68. **Nakajima, N., H. Ozeki, and Y. Shimura.** 1981. *In vitro* transcription of the *supB-E* tRNA operon of *Escherichia coli*: characterization of transcription products. J. Biol. Chem. **257**:11113–11120.
69. **Nierlich, D. P.** 1978. Regulation of bacterial growth, RNA and protein synthesis. Annu. Rev. Microbiol. **32**:393–432.
70. **Nishimura, S.** 1978. Modified nucleosides and isoaccepting tRNA, p. 168–195. *In* S. Altman (ed.), Transfer RNA. MIT Press, Cambridge, Mass.
71. **Nomura, M., R. Gourse, and G. Baughman.** 1984. Regulation of the synthesis of ribosomes and ribosomal components. Annu. Rev. Biochem. **53**:75–117.
72. **Ozeki, H.** 1980. The organization of transfer RNA genes in *Escherichia coli*, p. 173–183. *In* S. Osawa, H. Ozeki, H. Uchida, and T. Yura (ed.), Genetics and evolution of RNA polymerase, tRNA and ribosomes. University of Tokyo Press, Tokyo.
73. **Ozeki, H., H. Inokuchi, F. Yamao, M. Kodaira, H. Sakano, T. Ikemura, and Y. Shimura.** 1980. Genetics of nonsense suppressor tRNAs in *Escherichia coli*, p. 341–362. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
74. **Quay, S. C., and D. L. Oxender.** 1980. Role of $tRNA^{Leu}$ in branched-chain amino acid transport, p. 481–491. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
75. **Ray, B. K., and D. Apirion.** 1981. Transfer RNA precursors are accumulated in *Escherichia coli* in the absence of RNase E. Eur. J. Biochem. **114**:517–524.
76. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
77. **Rossi, J., J. Egan, L. Hudson, and A. Landy.** 1981. The *tyrT* locus: termination and processing of a complex transcript. Cell **26**:305–314.
78. **Rossi, J. J., and A. Landy.** 1979. Structure and organization of the two $tRNA^{Tyr}$ gene clusters on the *E. coli* chromosome. Cell **16**:523–534.
79. **Russell, R. L., J. N. Abelson, A. Landy, M. L. Gefter, S. Brenner, and J. D. Smith.** 1970. Duplicated genes for tyrosine transfer RNA in *E. coli*. J. Mol. Biol. **47**:1–13.
80. **Ryan, M. J., E. L. Brown, R. Belagaje, and H. G. Khorana.** 1980. Cloning of two chemically synthesized genes for a precursor to the $Su_3^+$ suppressor $tRNA^{Tyr}$, p. 245–258. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
81. **Sakano, H., and Y. Shimura.** 1978. Characterization and *in vitro* processing of transfer RNA precursors accumulated in a temperature-sensitive mutant of *Escherichia coli*. J. Mol. Biol. **123**:287–326.
82. **Sanderson, K. E., and J. R. Roth.** 1983. Linkage map of *Salmonella typhimurium*, edition VI. Microbiol. Rev. **47**: 410–453.
83. **Schwartz, I., R.-A. Klotsky, D. Elseviers, P. J. Gallagher, M. Krauskopf, M. A. Q. Siddiqui, J. F. H. Wong, and B. A. Roe.** 1983. Molecular cloning and sequencing of *pheU*, a gene for *Escherichia coli* $tRNA^{Phe}$. Nucleic Acids Res. **11**:4379–4389.
84. **Sekiya, T., R. Contreras, H. Küpper, A. Landy, and H. G. Khorana.** 1976. *Escherichia coli* tyrosine transfer ribonucleic acid genes: nucleotide sequences of their promoters and of the regions adjoining C-C-A ends. J. Biol. Chem. **251**:5124–5140.
85. **Sekiya, T., M. Mori, N. Takahashi, and S. Nishimura.** 1980. Sequence of the distal $tRNA_1^{Asp}$ gene and the transcription termination signal in the *Escherichia coli* ribosomal RNA operon *rrnF* (or *G*). Nucleic Acids Res. **8**:3809–3827.
86. **Sekiya, T., and S. Nishimura.** 1979. Sequence of the gene for isoleucine $tRNA_1$ and the surrounding region in a ribosomal RNA operon of *Escherichia coli*. Nucleic Acids Res. **6**:575–592.
87. **Shimura, Y., H. Sakano, S. Kubokawa, F. Nagawa, and H. Ozeki.** 1980. tRNA precursors in RNase P mutants, p. 43–58. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
88. **Skjold, A. C., H. Juarez, and C. Hedgcoth.** 1973. Relationships among deoxyribonucleic acid, ribonucleic acid, and specific transfer ribonucleic acids in *Escherichia coli* $15T^-$ at various growth rates. J. Bacteriol. **115**:177–187.
89. **Soffer, R. L.** 1980. Biochemistry and biology of aminoacyl-tRNA-protein transferases, p. 493–505. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
90. **Sprinzl, M., and D. H. Gauss.** 1984. Compilation of tRNA sequences. Nucleic Acids Res. **12**(Suppl.):r1–r57.
91. **Sprinzl, M., and D. H. Gauss.** 1984. Compilation of sequences of tRNA genes. Nucleic Acids Res. **12**(Suppl.):r59–r131.
92. **Steege, D. A.** 1983. A nucleotide change in the anticodon of an *E. coli* serine transfer RNA results in $supD^-$ amber supression. Nucleic Acids Res. **11**:3823–3832.
93. **Steege, D. A., and J. I. Horabin.** 1983. Temperature-inducible amber suppressor: construction of plasmids containing the *Escherichia coli* $serU^-$ ($supD^-$) gene under control of the bacteriophage lambda $p_L$ promoter. J. Bacteriol. **155**:1417–1425.
94. **Tamura, F., S. Nishimura, and M. Ohki.** 1984. The *E. coli divE* mutation, which differentially inhibits synthesis of certain proteins, is in $tRNA_1^{Ser}$. EMBO J. **3**:1103–1107.
95. **Thorbjarnardóttir, S., T. Dingermann, T. Rafnar, O. S. Andrésson, D. Söll, and G. Eggertsson.** 1985. Leucine tRNA family of *Escherichia coli*: nucleotide sequence of the *supP*(Am) suppressor gene. J. Bacteriol. **161**:219–222.
96. **Thorbjarnardóttir, S., H. Uemura, T. Dingermann, T. Rafnar, S. Thorsteinsdóttir, D. Söll, and G. Eggertsson.** 1985. *Escherichia coli supH* suppressor: temperature-sensitive missense suppression caused by an anticodon change in $tRNA_2^{Ser}$. J. Bacteriol. **161**:207–211.

97. **Travers, A. A.** 1980. Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. J. Bacteriol. **141:**973–976.
98. **Travers, A. A.** 1984. Conserved features of coordinately regulated *E. coli* promoters. Nucleic Acids Res. **12:**2605–2618.
99. **Travers, A. A., A. I. Lamond, and H. A. F. Mace.** 1982. ppGpp regulates the binding of two RNA polymerase molecules to the *tyrT* promoter. Nucleic Acids Res. **10:**5043–5057.
100. **Travers, A. A., A. I. Lamond, H. A. F. Mace, and M. L. Berman.** 1983. RNA polymerase interactions with the upstream region of the *E. coli tyrT* promoter. Cell **35:**265–273.
101. **Tucker, S. D., A. S. Gopalakrishnan, R. Bollinger, W. Dowhan, and E. J. Murgola.** 1982. Molecular mapping of *glyW*, a duplicated gene for $tRNA_3^{Gly}$ of *Escherichia coli*. J. Bacteriol. **152:**773–779.
102. **Umbarger, H. E.** 1980. Comments on the role of aminoacyl-tRNA in the regulation of amino acid biosynthesis, p. 453–467. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
103. **Vold, B. S.** 1985. Structure and organization of genes for transfer RNA in *Bacillus subtilis*. Microbiol. Rev. **49:**71–80.
104. **VonBogelen, R. A., V. Vaughe, and F. C. Neidhardt.** 1983. Gene for heat-inducible lysyl-tRNA synthetase (*lysU*) maps near *cadA* in *Escherichia coli*. J. Bacteriol. **153:**1066–1068.
105. **von Hippel, P. H., D. G. Bear, W. D. Morgan, and J. A. McSwiggen.** 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. Annu. Rev. Biochem. **53:**389–446.
106. **Weissbach, H., and S. Pestka (ed.).** 1977. Molecular mechanisms of protein biosynthesis. Academic Press, Inc., New York.
107. **Wu, H., and D. M. Crothers.** 1984. The locus of sequence-directed and protein-induced DNA bending. Nature (London) **308d:**509–513.
108. **Yoshimura, M., H. Inokuchi, and H. Ozeki.** 1984. Identification of transfer RNA suppressor in *Escherichia coli*. IV. Amber suppressor $Su^+6$, a double mutant of a new species of leucine tRNA. J. Mol. Biol. **177:**627–644.
109. **Yoshimura, M., M. Kimura, M. Ohno, H. Inokuchi, and H. Ozeki.** 1984. Identification of transfer RNA suppressors in *Escherichia coli*.III. Ochre suppressors of lysine tRNA. J. Mol. Biol. **177:**609–625.
110. **Young, R. A.** 1979. Transcription termination in the *Escherichia coli* ribosomal RNA operon *rrnC*. J. Biol. Chem. **254:**12725–12731.
111. **Young, R. A., R. J. Bram, and J. A. Steitz.** 1980. rRNA and tRNA processing signals in the rRNA operons of *Escherichia coli*, p. 99–106. *In* D. Söll, J. N. Abelson, and P. R. Schimmel (ed.), Transfer RNA: biological aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
112. **Young, R. A., R. Macklis, and J. A. Steiz.** 1979. Sequence of the 16S-23S spacer region in two ribosomal operons of *Escherichia coli*. J. Biol. Chem. **254:**3264–3271.